

| Sequence          | Strd Orig | zScore  | EScore  | Len      | Documentation                            |
|-------------------|-----------|---------|---------|----------|--|
| gb.est71:BE420587 | +         | 1038.50 | 2078.33 | 9.4e-107 | 1212   BE420587 HWMO00.D12 ITEC HWM      |
| gb.est80:C99909   | +         | 1028.00 | 2072.77 | 1.9e-106 | 651   C99909 C99909 Atabidopsis thaliana |
| gb.est74:BE465820 | +         | 714.50  | 1432.96 | 8.3e-71  | 653   BE465820 SP43E04NR1F1000 Nodul     |
| gb.est71:BE440835 | +         | 694.00  | 1392.38 | 1.2e-68  | 502   BE440835 SP42F03.Y1 Gm-c1043 GL    |
| gb.est37:AI938253 | +         | 684.00  | 1374.37 | 1.5e-67  | 486   AI938253 SC41h02.Y1 Gm-c1014 GL    |
| gb.est73:BE643915 | +         | 632.00  | 1268.69 | 1.2e-61  | 469   BE643915 NCX1.048.H12_F NCX1 (N    |
| gb.est76:BE803390 | +         | 590.50  | 1181.50 | 8.5e-57  | 574   BE803390 sr55G005.Y1 Gm-c1051 GL   |
| gb.est36:AV408114 | +         | 582.00  | 1167.91 | 4.8e-56  | 424   AV408114 AV408114 Lotus japonicus  |
| gb.est71:BE433982 | +         | 554.50  | 1109.28 | 8.9e-53  | 519   BE433982 AV4081050 tomato break    |
| gb.est73:BE607262 | +         | 544.00  | 1091.05 | 9.2e-52  | 401   BE607262 NC132_G06_F NCX1 (N       |
| gb.est80:C91832   | +         | 522.00  | 1046.88 | 2.7e-49  | 378   C91832 C91832 Rice panicle short   |
| gb.est48:AW694621 | +         | 496.50  | 986.24  | 1.3e-46  | 757   AW694621 W078C11ST1F1085 devel     |
| gb.est70:BE403631 | +         | 492.50  | 980.71  | 1.3e-45  | 612   BE403631 WHE0434.H04_O0825 Whea    |
| gb.est73:BE600656 | +         | 463.00  | 921.70  | 2.5e-42  | 556   BE600656 P11_91.A10.B1.A002 Pat    |
| gb.est27:AI941126 | +         | 442.00  | 884.51  | 3.0e-40  | 352   AI941126 SB84h02.Y1 Gm-c1010 GL    |
| gb.est48:AW691011 | +         | 428.50  | 849.29  | 2.7e-38  | 654   AW691011 W040D08ST1F1000 devel     |
| gb.est86:7F5860   | +         | 415.00  | 830.21  | 3.1e-37  | 330   7F5860 10638 Lambda-PRU2 Arabid    |
| gb.est46:AW562918 | +         | 414.50  | 823.39  | 7.5e-37  | 527   AW562918 6019C11.Y1 Gm-c1031 GL    |
| gb.est72:AI993308 | +         | 395.50  | 785.32  | 9.9e-35  | 498   AW562760 sb19C11.Y1 Gm-c1031 GL    |
| gb.est37:AI933344 | +         | 381.00  | 749.28  | 1.0e-32  | 838   AI933344 660069F09.X1 660 - Mix    |
| gb.est72:BE498756 | +         | 380.00  | 754.56  | 5.1e-33  | 464   BE498756 WHE0965.C04_F0725 Whea    |
| gb.est74:BE704796 | +         | 378.50  | 746.41  | 1.5e-32  | 700   BE704796 SC02_02F04.A SC02.AAPD    |
| gb.est73:BE597622 | +         | 358.50  | 707.69  | 1.2e-30  | 591   BE597622 P11_72.A11.B1.A002 Pat    |
| gb.g8s25:AG021043 | +         | 334.50  | 657.90  | 1.2e-27  | 631   AG021043 Oryza sativa DNA, 3' /    |
| gb.est73:BE597324 | +         | 292.50  | 570.94  | 8.6e-23  | 698   BE597324 P11_72.A11.GL.A002 Pat    |
| gb.est36:AV383444 | +         | 289.50  | 569.17  | 1.1e-22  | 491   AV383444 AV383444 Halocynthia r    |
| gb.est69:BE321087 | +         | 285.00  | 553.00  | 8.6e-22  | 863   BE321087 NF033.H11RT1F1092 devel   |
| gb.est39:AW062071 | -         | 278.50  | 544.88  | 2.4e-21  | 570   AW062071 660009C11.Y1 660 - Mix    |
| gb.g8s13:AO916553 | +         | 278.00  | 540.18  | 4.5e-21  | 767   AO916553 nb8e005020F CUGI Rice     |
| gb.est48:AW720517 | +         | 277.00  | 542.03  | 3.5e-21  | 560   AW720517 nb8e013e3r Lotus japon    |
| gb.est45:AW498397 | +         | 266.50  | 519.79  | 6.1e-20  | 598   AW498397 660047D08.X1 660 - Mix    |
| gb.g8s10:AO687325 | +         | 257.50  | 500.53  | 7.2e-19  | 643   AO687325 sb5b0074M20R CUGI Rice    |
| gb.est46:AW566326 | +         | 257.00  | 513.55  | 4.1e-19  | 207   AW566326 660069F09.Y1 660 - Mix    |
| gb.est67:BE131390 | +         | 250.50  | 487.33  | 3.9e-18  | 589   BE131390 L48-1365T3 Ice plant      |
| gb.est67:BE131520 | +         | 247.50  | 482.94  | 6.9e-18  | 512   BE131520 L48-1504T3 Ice plant      |
| gb.est37:AI938086 | +         | 246.50  | 481.47  | 8.3e-18  | 489   AI938086 SC41h08.X1 Gm-c1014 GL    |
| gb.est43:AW368451 | +         | 246.50  | 477.89  | 1.3e-17  | 653   AW368451 CM4 -HT0193-061099-022    |
| gb.est29:AW094795 | +         | 246.00  | 475.81  | 1.7e-17  | 711   AW094795 AO949795 Rice panicle     |
| gb.est48:AW053136 | +         | 237.50  | 464.11  | 7.7e-17  | 451   AW053136 sk59d08.Y1 Gm-c1019 GL    |
| gb.est45:AW058702 | +         | 235.00  | 457.78  | 1.7e-16  | 498   AW058702 S135d06.Y1 Gm-r1030 GL    |
| gb.est41:AW191191 | -         | 229.50  | 448.28  | 5.9e-16  | 433   AW191191 687016F08.X1 687 - Rat    |

```
|||||.....:|||||.....:|||||.....:|||||.....:|||||.....:
207 GGGCAATATTCGCGAAGCGACTCTCTCGCACCTTGCAGGTGTAAAG 256
166 nGlnAspGluThrProLeuLeuTyrSerLeuValPheGlyGluGlyVal 183
:|||||.....:|||||.....:|||||.....:|||||.....:|||||.....:
257 CCAAGATGAGACACTTCTTGACATTTGGTGGTGGTGGAGGGTTG 306
183 aAsnAspAlaThrSerValValPheAsnAlaIleGlnSerPheAsp 199
|||||.....:|||||.....:|||||.....:|||||.....:|||||.....:
307 TTAACGATGGACATCAGTTGTTGTTGTTCAATGCAATCCAGAACTTGAT 356
200 LeuThrHisLeuAsnHisGluAlaAlaPheHisLeuLeuGlyAsn 216
||| :|||||.....:|||||.....:|||||.....:|||||.....:
357 CTTCGAATTCAGTAGCCCTCAAAATTCACAAATTCATTGGAATTCCT 406
216 uTyrLeuPheLeuLeuSerThrLeuLeuGlyAlaAlaThrGlyLeuIle 233
|||||.....:|||||.....:|||||.....:|||||.....:|||||.....:
407 CTATCTATTGGGCCCGCCAGTACCTTCTGGAGTATCTAGTGAGACTCTCA 456
233 eAlaTyrValIleLysLysLeuTyrPheGlyArgHisSerThrAspArg 249
457 GTGCTTATGTCATCAAGAACTGACTTTGGCAGCACTCCACTGATCGT 506
250 GluValAlaLeuMetMetLeuMetAlaTyrLeuSerTyrMetLeuAlaG1 266
|||||.....:|||||.....:|||||.....:|||||.....:|||||.....:
507 GAAGTTCGTATTATGATGCTCATGGCTTATTATCTTACATGCTGGCTGA 556
266 uLeuPheAspLeuSerGlyIleLeuThrValPhePheCysGlyIleValM 283
557 ATTGCTTTGATTGAGTGGTATCTTCACGGTTTCTTCTGGTATTGTAA 606
283 eTyrHisTyrThrTyrHisAsnValThrGluSerSerArgIleThrThr 299
607 TGTGCACTATACCTGGCACAATGTACACAGAGTTCACGGGTCACAACC 656
300 LysHisThrPheAlaThrLeuSerPheLeuAlaGluThrPheIlePheLe 316
|||||.....:|||||.....:|||||.....:|||||.....:|||||.....:
657 AAGCATGCTTTGGCCACATGCTATTCATCTCTGAGACGTTCTCTTTCT 706
316 uTyrValGlyMetAspAlaLeuAspIleAspLysTyrArgSerValSerA 333
707 CTATGTTGGCATGATGATTCATGGATATAGAGAGTGGAAAATTGGTAGTG 755
333 spthr.....ProGlyThrSerIleAlaValSerSerIleLeuMetGly 347
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
756 A ACATATAGCCCAATCAATCTATTGCTTGGCTGAGCTCCATTATTTGGCG 805
348 LeuValMetValGlyArgAlaAlaPheValPheProLeuSerPheLeuSe 364
|||||.....:|||||.....:|||||.....:|||||.....:|||||.....:
806 TTGCTGTGGTTGCAAGAGCTGCATTTGTTTCCCTCTATCT...ATCTC 852
364 r.AsnLeuAlaLysLysAsnGlnSerGluLysIleAsnPheAsnMetGln 380
| :|||||.....:|||||.....:|||||.....:|||||.....:|||||.....:
853 TCCAAATGACCAAAAACCTCCAGCGGAGAGATCTCTGTAGGCAGCAA 902
381 ValValIleThrPyrSerGlyLeuMetArgGlyAlaValSerMet 395
|||||.....:|||||.....:|||||.....:|||||.....:|||||.....:
903 GTTATTATTTGGTGGCG.GGGCTCATGAGAGGGGGCGGTGTCAATG 946
```

seq\_name: gb\_est80:C99909

```
seq_documentation_block: 651 bp mRNA EST 21-SEP-1999
LOCUS C99909
DEFINITION C99909 Arabidopsis thaliana library (Motohashi R) Arabidopsis
thaliana cDNA clone 65, mRNA sequence.
ACCESSION C99909
VERSION C99909.1 GI:4714029
KEYWORDS EST.
SOURCE thale cress.
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
Brassicales; Brassicaceae; Arabidopsis.
```

```
REFERENCE 1 (bases 1 to 651)
AUTHORS Motohashi,R., Shinozaki,K. and Yamaguchi-Shinozaki,K.
TITLE Arabidopsis thaliana YAC C1C3B1 region specific cDNA,clone 2
JOURNAL Unpublished (1999)
COMMENT Contact: Motohashi R
Laboratory of Plant Molecular Biology
Institute of Physical and Chemical Research(RIKEN)
3-1-1, Koyadai, Tsukuba 305-0074, Japan
Tel: +81-298-36-4359
Email: motohashi@tc.riken.go.jp.
FEATURES
source
1..651
/organism="Arabidopsis thaliana"
/db_xref="taxon:3702"
/clone="65"
/clone_lib="Arabidopsis thaliana library (Motohashi R)"
BASE COUNT 161 a 156 c 152 g 178 t 4 others
ORIGIN
```

```
alignment_scores:
Quality: 1028.00 Length: 220
Ratio: 4.849 Gaps: 4
Percent Similarity: 96.364 Percent Identity: 95.909
alignment_block:
US-09-271-584A-2 x C99909 ..
```

Align seg 1/1 to: C99909 from: 1 to: 651

```
314 IlePheLeuTyrValGlyMetAspAlaLeuAspIleAspLysTyrArgSe 330
|||||.....:|||||.....:|||||.....:|||||.....:|||||.....:
1 ATTTCCTGATGTTGGAATGATGCTTGGACATTGACAAATGGAGATC 50
330 rValSerAspThrProGlyThrSerIleAlaValSerSerIleLeuMetG 347
|||||.....:|||||.....:|||||.....:|||||.....:|||||.....:
51 CGTGTGATGACACACCGGGAACATCGATCGAGTGCATCAATCTTAATGG 100
347 lLeuValMetValGlyArgAlaAlaPheValPheProLeuSerPheLeu 363
|||||.....:|||||.....:|||||.....:|||||.....:|||||.....:
101 GTCGTGTCATGTTGGAAGAGCAGCGTTCGTCCTTTCCGTTATCGTTCTA 150
364 SerAsnLeuAlaLysLysAsnGlnSerGluLysIleAsnPheAsnMetG1 380
|||||.....:|||||.....:|||||.....:|||||.....:|||||.....:
151 TCTAACTTAGCCCAAGAAGAAATCAAAGCGAGAAAATCAACTTTAATGCA 200
380 nValValIleThrTyrSerGlyLeuMetArgGlyAlaValSerMetAlaL 397
|||||.....:|||||.....:|||||.....:|||||.....:|||||.....:
201 GGTGTGATTTGGTGGTCTGGTCTCATGAGAGTGTGTATCTATGGCTC 250
397 euAlaTyrAsnLysPheThrArgAlaGlyHisThrAspValArgGlyAsn 413
|||||.....:|||||.....:|||||.....:|||||.....:|||||.....:
251 TTGCATACACAAGTTTACAAGCGCGGCGCACAGATGTAGCGGGAAAT 300
414 AlaIleMetIleThrSerThrIleThrValCysLeuPheSerThrValVa 430
|||||.....:|||||.....:|||||.....:|||||.....:|||||.....:
301 GCAATCATGATCAGCAGTACGATAACTGTCTGCTCTTTTAGCACAGTGT 350
430 lPheGlyMetLeuThrLysProLeuIleSerTyrLeuLeuProHisGlnA 447
|||||.....:|||||.....:|||||.....:|||||.....:|||||.....:
351 GTTGTGATGTGACCAACCACTCATAGCTACCTATTACCGCCACAGA 400
447 snAlaThrThrSerMetLeuSerAspAsnThrProLysSerIleHis 463
|||||.....:|||||.....:|||||.....:|||||.....:|||||.....:
401 AGCCACCAACAGCATGTTATCTGATGACACACACCCCAAAATCCATACAT 450
464 lIleProLeuLeuAspGlnAspSerPheIleGluProSerGlyAsnHisAs 480
|||||.....:|||||.....:|||||.....:|||||.....:|||||.....:
451 ATCCCTTTTGTGGACCAAGACCG.TTCATTGAGCCTTCAGGGAACACAA 499
480 nValProArgProAspSerIleArgGlyPheLeuThrArgProThrArgT 497
|||||.....:|||||.....:|||||.....:|||||.....:|||||.....:
500 TGTGCTCGGCTGACAGTATACGTGCTTCTTGACACGGGCCCTCSCAA 549
```

497 hrValHisTyrTrp.ArgGlnPheAspSerPheMetArgProVa 513  
 |||||  
 550 CCGTGATTACTACTGGGAGCAATTTGATGACTCTTCCGCGACCGGT 599  
 |||||  
 513 lPhe.GlyGlyArgGlyPheValProPheVal.Pro.GlySerProThrG 529  
 |||||  
 600 CTTGGGAGGTGGTCTTTGACCTTTGTTTCCARGGTTCTCCAACTG 649  
 529 lu 529  
 ||  
 650 AR 651

seq\_name: gb\_est47:AW685820

seq\_documentation\_block: 653 bp mRNA EST 15-JUN-2000  
 LOCUS AW685820 NF035804NR1F1000 Nodulated root Medicago truncatula cDNA clone  
 DEFINITION NF035804NR 5', mRNA sequence.

ACCESSION AW685820 GI:7560556

VERSION EST.

KEYWORDS barrel medic.

SOURCE Medicago truncatula

ORGANISM Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;  
 Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I;  
 Fabales; Fabaceae; Papilionoideae; Medicago.

REFERENCE 1 (bases 1 to 653)

AUTHORS Watson,B.S., Shin,H.-S., Lopez-Meyer,M., Scott,A.D., Harris,A.R.,  
 Gonzales,R.A., Bell,C.J., Flores,H.R., Iman,J.T., Weller,J.W., May  
 ,G.D. and Paiva,N.L.

TITLE Expressed Sequence Tags from the Samuel Roberts Noble Foundation

JOURNAL Medicago truncatula nodulated root library

COMMENT Unpublished (2000)

Contact: Paiva NL

Plant Biology Division

The Samuel Roberts Noble Foundation

2510 Sam Noble Parkway, Ardmore, OK 73402, USA

Tel: 580 221 7317

Fax: 580 221 7380

Email: nlpaiva@noble.org

Insert Length: 653 Std Error: 0.00

Plate: 035 row: E column: 04

Seq primer: TCACACAGGAACAGCTATGAC.

FEATURES

source

1..653

/organism="Medicago truncatula"

/db\_xref="taxon:3880"

/clone="NF035804NR"

/tissue\_type="root"

/dev\_stage="Pooled developmental"

/note="Vector: Lambda Zap; Four-week old Rhizobium

meliloti-inoculated Medicago truncatula roots, containing

a mixture of young and old roots and nodules."

BASE COUNT 167 a 133 c 148 g 205 t

ORIGIN

alignment\_scores:

Quality: 714.50 Length: 218

Ratio: 3.841 Gaps: 2

Percent Similarity: 85.321 Percent Identity: 64.220

alignment\_block:

US-09-271-584A-2 x AW685820 ..

Align seg 1/1 to: AW685820 from: 1 to: 653

307 serPheLeuAlaGluThrPheLeuTyrValGlyMetAspAlaLe 323

|||||

2 TCATTGCTGAGATATATCTTCTTATGTTGGGATGGATGCAC 51

323 uAspIleAspLysTyrArgSerValSerAspThrProGlyThrSerIleA 340  
 |||||  
 52 ACATATGGAGAAAGTGGCATTCGTAAGTCAAGGCCAAAAAATCAATTG 101  
 |||||  
 340 laValSerSerIleLeuMetGlyLeuValMetValGlyArgAlaAlaPhe 356  
 ::|||  
 102 GGGTCAGTTTCATTGCTGTGGCAGCTATCTAGTGGGAAGAGCTGCATTT 151  
 |||||  
 357 ValPheProLeuSerPheLeuSerAsnLeuAlaLysLysAsnGlnSerGl 373  
 |||||  
 152 GTTTCCTCTCTTCTTCATATCCAACTGTCTAAGAAGTCTCAATCTGA 201  
 |||||  
 373 ulysIleAsnPheAsnMetGlnValValIleTrpSerGlyLeuMeta 390  
 |||||  
 202 GAAATTTAGTTTAAAGCAACAAAGTGAATTTGGTGGCGCTGCTCATGC 251  
 |||||  
 390 rGClValAlaValSerMetAlaLeuAlaTyrAsnLysPheThrArgAlaGly 406  
 ::|||  
 252 GTGGAGCTGTTCTATGTGCACTGTCTTACAAATGATGACAAAGCCACTAGGA 301  
 |||||  
 407 HisThrAspValArgGlyAsnAlaIleMetIleThrSerThrIleThrVa 423  
 |||||  
 302 CATACTAAATTCGCTGAGAAAGTGAATTCATGATCACTAGTACTATTACAGT 351  
 |||||  
 423 lCysLeuPheSerThrValValPheGlyMetLeuThrLysProLeuIle 440  
 |||||  
 352 TGTACTCTTTAGCAGCTGTGTTGGGTTGATGACAAAGCCACTTGTGA 401  
 |||||  
 440 eTyrLeuLeuPro...HisGlnAsnAlaThrThrSerMetLeuSerAsp 455  
 |||||  
 402 GGTATTGCTACCTCTTCTTAAACACATAATAATCAGCATACCGTCCCG 451  
 |||||  
 456 AspAsnThrProLysSerIleHisIleProLeuLeuAspGlnAspSerPh 472  
 |||||  
 452 CCATCTAGCCCGAAATCATTCAGTGTGCACTTCTTGGCAATGGAGAA... 499  
 |||||  
 472 eIleGluProSerGlyAsnHisAsnValProArgProAspSerIleArgG 489  
 ::|||  
 500 ....GATGTAGGGGCAATGGTGGTACCCAAAGACCAAGCAGATTCGCTA 545  
 |||||  
 489 lypheLeuThrArgProThrArgThrValHisTyrTyrTrpArgGlnPhe 505  
 ::|||  
 546 CGCTCCTTAGGATTCCTAGTATGCGGTGATCATCACTACTGCGCAAAATTT 595  
 |||||  
 506 AspAspSerPheMetArgProValPheGlyGlyArgGlyPheValProPh 522  
 |||||  
 596 GATGATTCCTTCATGCGACCTGCTTTGGTGGCGGGGCTTTTGTACCTTA 645  
 |||||

seq\_name: gb\_est71:BE440835

seq\_documentation\_block:

LOCUS BE440835

DEFINITION sp42f03.y1 Gm-cl043 Glycine max cDNA clone GENOME SYSTEMS CLONE ID:

Gm-cl043-918 5' similar to TR:004655 004655 SIMILAR TO

SODIUM/HYDROGEN EXCHANGER. ; mRNA sequence.

ACCESSION BE440835

VERSION BE440835.1 GI:9440328

KEYWORDS EST.

SOURCE soybean.

ORGANISM Glycine max

Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;

Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I;

Fabales; Fabaceae; Papilionoideae; Glycine.

REFERENCE 1 (bases 1 to 502)

AUTHORS Shoemaker,R., Keim,P., Vodkin,L., Erpelting,J., Corvill,V., Khanna

,A., Holla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C.,

Wyllie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers

,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk

,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann

25-JUL-2000

```
R., Waterston, R. and Wilson, R.
Public Soybean EST Project
Unpublished (1999)
Contact: Shoemaker R/Public Soybean EST Project
Public Soybean EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available through: Genome Systems, Inc. 4633 World
Parkway Circle St. Louis, Missouri 63134 For further information
call: (800) 430-0030 or (314) 427-3222 FAX: (888) 919-3324 or (314)
427-3324 or contact: clones@genomesystems.com or
info@genomesystems.com web site: www.genomesystems.com
High quality sequence stop: 470.
Location/Qualifiers
1. 502
/organism="Glycine max"
/db_xref="taxon:3847"
/clone="GENOME SYSTEMS CLONE ID: Gm-cl043-918"
/clone_lib="Gm-cl043"
/tissue_type="Hypocotyl and plumule, germinating seeds"
/lab_host="DH10B"
/note="Vector: pT73Pac (Pharmacia); Site 1: EcoRI;
Site 2: NotI; This cDNA library was constructed from mRNA
isolated from hypocotyl and plumule tissues of seeds
germinated for three days of the cultivar Williams.
Complementary DNA was synthesized from mRNA using a primer
consisting of a poly(dT) sequence with a NotI restriction
site. EcoRI adapters were ligated to the blunt-ended cDNA
fragments followed by digestion with EcoRI and NotI. The
cDNA fragments were directionally cloned into the
EcoRI-NotI restriction site of the pT73-Pac vector. The
ligated cDNA fragments were transformed into DH10B host
cells (Gibco BRL). This library was constructed by Dr.
Randy Shoemaker."
BASE COUNT 119 a 95 c 112 g 176 t
ORIGIN

alignment_scores:
Quality: 694.00 Length: 167
Ratio: 4.365 Gaps: 0
Percent Similarity: 95.210 Percent Identity: 81.437

alignment_block:
US-09-271-584A-2 x BE440835 ..
Align seg 1/1 to: BE440835 from: 1 to: 502

112 LeuPheGlyAlaValGlyThrIleLeSerCysThrIleLeSerLeuG1 128
|||||
1 TTGTTTGTGTGTTATGGCACAATTAATATCTTGTACCAATCAATCTTTGGG 50
|||||
128 yValThrGlnPhePheLysLysLeuAspIleGlyThrPheAspLeuGlyA 145
|||||
51 TCCACACAAATTTTAAGAGGTGGATGTTGGTCCCACTGGAATTAGGGG 100
|||||
145 sPtyrLeuAlaIleGlyAlaIlePheAlaAlaThrAspSerValCysThr 161
|||||
101 ATTCTCTAGCAATTGGTCAATATTGCTGCAACGGATTCTGTTGGACA 150
|||||
162 LeuGlnValLeuAsnGlnAspGluThrProLeuLeuTyrrSerLeuValP 178
|||||
151 TTGCAGGTGCTAAATCAGATGAGACACCTTGTGTTACATCTTGATT 200
|||||
178 eGlyGluGlyValValAsnAspAlaThrSerValValValPheAsnAlaI 195
|||||
201 TGGGGAGGGTGTGTTGAATGATGCTACATCAGTGGTGGCTTTTCATGCA 250
|||||
195 leGlnSerPheAspLeuThrHisLeuAsnHisGluAlaAlaPheHisLeu 211
|||||
```



plants. Complementary DNA was synthesized from mRNA using a 3' anchored poly (dT) primer. EcoRI adapters were ligated to the blunt-ended cDNA fragments followed by digestion with EcoRI and HindIII. The cDNA fragments were directionally cloned into the EcoRI-HindIII restriction site of the p773-Pac vector. The ligated cDNA fragments were transformed into DH10B host cells (Gibco BRL). This library was constructed by Dr. Randy Shoemaker and Dr. John Erpelidg.

BASE COUNT 122 a 101 c 99 g 164 t  
ORIGIN

## alignment\_scores:

Quality: 684.00 Length: 160  
Ratio: 4.442 Gaps: 0  
Percent Similarity: 96.250 Percent Identity: 79.375

## alignment\_block:

US-09-271-584A-2 x A1938253 ..

Align seg 1/1 to: A1938253 from: 1 to: 486

274 LeuThrValPhePheCysGlyIleValMetSerHisThrTrpHisAs 290

|||||  
7 CTCACCTGTATCTTTTGTGGATTGTCTCATCTCATTTACCTGCATAA 56

290 nValThrGluSerSerArgIleThrThrLysHisThrPheAlaThrLeuS 307

|||||  
57 TGTGACTGAGAGTTCAAGATCACCACCAAGCATCTCTTTTGCACCTTGT 106

307 erPheLeuAlaGluThrPheIlePheLeuTyrValGlyMetAspAlaLeu 323

|||||  
107 CTTTGTCTCAGATCTTTATCTTCCTTTAGTTGGTATGATGCCCTTG 156

324 AspIleAspLysTrpArgSerValSerAspThrProGlyThrSerIleAl 340

|||||  
157 GACATTGAAAAATGGAATTTTCTCAGTGATAGCCCTGGAACATCTGTAGC 206

340 aValSerSerIleLeuMetGlyLeuValMetValGlyArgAlaAlaPheV 357

|||||  
207 ACTAGTTCAGTATTGTTGGGCTTAATCTTCTTGGGAAGAGCAGCTTTTG 256

357 alPheProLeuSerPheLeuSerAsnLeuAlaLysLysAsnGlnSerGlu 373

|||||  
257 TTTTCCCTTATCCTTCTATCCACTTGGCTAAAAATCACCATAATGAG 306

374 LysIleAsnPheAsnMetGlnValValIleTrpTrpSerGlyLeuMetAr 390

|||||  
307 AAAATCAGCTTCAGACAGCAAGATTATCATTTTGGTGGCTGCTTATGAG 356

390 gGlyAlaValSerMetAlaLeuAlaTyrAsnLysPheThrArgAlaGlyH 407

|||||  
357 AGGTCTGTTTCAATCGCAGCTTGGCATATAATCAGTTCCACCATGTGGCC 406

407 IsThrAspValArgGlyAsnAlaIleMetIleThrSerThrIleThrVal 423

|||||  
407 ACACCTTCACTAAGAAGCAATCAATCATGATCAACAGCACCATCAGTCT 456

424 CysLeuPheSerThrValValPheGlyMet 433

|||||  
457 GTGCTTTTCAGCACAGCAGTGGTGTGGTCTG 486

seq\_name: gb\_est73:BE643915

seq\_documentation\_block:

LOCUS BE643915 469 bp mRNA EST  
DEFINITION NCX1\_048\_H12\_F NXCI (Nsf Xylem Compression wood Inclined) Pinus  
taeda cDNA clone NXCI\_048\_H12 5', mRNA sequence.

ACCESSION BE643915

VERSION BE643915.1 GI:9956522

KEYWORDS EST.

SOURCE loblolly pine.

## ORGANISM

Pinus taeda

Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Pinus; Pinus.

## REFERENCE

1 (bases 1 to 469)

## AUTHORS

Sederoff, R.

## TITLE

Molecular Basis of Wood Formation in the Pine Megagenome

## JOURNAL

Unpublished (2000)

## COMMENT

Contact: Johnson, Arthur

North Carolina State University

Tel: 919 515 7800

Fax: 919 515 7801

Email: ajohnson@unity.ncsu.edu

Seq primer: T3.

## FEATURES

Location/Qualifiers

1..469

/organism="Pinus taeda"

/strain="Coastal plain loblolly pine from North Carolina"

/db\_xref="taxon:352"

/clone="NXCI\_048\_H12"

/tissue\_type="Xylem"

/cell\_type="Compression"

/lab\_host="Xl1-Blue"

/note="Vector: BlueScript SK; Site\_1: Eco RI; Site\_2: XhoI

; The library is from early (spring) wood, taken from

three six-year old trees (three different genotypes), in

the juvenile phase. These trees were induced to form

compression wood by bending to a 45 degree angle and tying

them to the ground. Differentiating xylem was harvested

from the bottoms of the inclined stems, and a mixture of

all three genotypes was used for the library. oligo-dT

primed cDNA was directionally cloned into the EcoRI-XhoI

BlueScript SK vector arms. NOTE: The sequences contain a

'cDNA adapter' between the EcoRI site and the start of the

EST. The adapter sequence is 'AATCGGCACGAG'.

EST. 109 a 91 c 97 g 155 t 17 others

## BASE COUNT

ORIGIN

## alignment\_scores:

Quality: 632.00 Length: 156

Ratio: 4.514 Gaps: 0

Percent Similarity: 89.744 Percent Identity: 77.564

## alignment\_block:

US-09-271-584A-2 x BE643915 ..

Align seg 1/1 to: BE643915 from: 1 to: 469

213 GlyAsnPheLeuTyrLeuPheLeuSerThrLeuGlyAlaAlaLath 229

|||||  
2 GGTAACTTTTGTTCATCTCTTTTGCACAAAGCACTGCATTGGGGATTGCAGC 51

229 rGlyLeuIleSerAlaTyrValIleLysLysLeuTyrPheGlyArgHis 246

|||||  
52 TGGCTTTTAAAGTCATCATCATCAAAAGCTTTGTACTTTTGGGAAGGCAC 101

246 erThrAspArgGluValAlaLeuMetMetLeuMetAlaTyrLeuSerTyr 262

|||||  
102 CTACTGATCGTGAAGTAGCAATTTATGATCATCATGCGGTATCTCTATAC 151

263 MetLeuAlaGluLeuPheAspLeuSerGlyIleLeuThrValPhePheCy 279

|||||  
152 ATGCTCGCTGAGCTGTTTTCACCTCAGTGGGATCTTGACTGTTTCTCTCTG 201

279 sGlyIleValMetSerHisThrThrPheHisAsnValThrGluSerSera 296

|||||  
202 TGCATTGTTATGCTCTACCTACACTTGGCATATATGACTGAAAGCTCAA 251

296 rgIleThrThrLysHisThrPheAlaThrLeuSerPheLeuAlaGluThr 312

|||||  
252 GGATTACCACCAAGCATGCAATTTGCAACAATGTCATTTATCGCTGAACA 301



SOURCE Lotus japonicus.  
ORGANISM Lotus japonicus  
Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids 1; Fabales; Fabaceae; Papilionoideae; Lotus.

REFERENCE 1 (bases 1 to 424)  
AUTHORS Asamizu, E., Nakamura, Y., Sato, S. and Tabata, S.  
TITLE Generation of 7137 non-redundant expressed sequence tags from a legume, Lotus japonicus  
JOURNAL DNA Res. 7 (2), 127-130 (2000)  
MEDLINE 20277479  
COMMENT Contact: Yasukazu Nakamura  
The First Laboratory for Plant Gene Research  
Kazusa DNA Research Institute  
Yana 1532-3, Kisarazu, Chiba 292-0812, Japan  
Email: ynakamu@kazusa.or.jp, URL: http://www.kazusa.or.jp/en/plant/.

FEATURES  
source  
1..424  
/organism="Lotus japonicus"  
/db\_xref="taxon:34305"  
/clone="MWL036dl1\_r"  
/clone\_lib="Lotus japonicus young plants (two-week old)"  
/dev\_stage="young plants (two-week old)"  
/note="vector: pBluescriptII SK-; Site\_1: EcoRI; Site\_2: XhoI; isolate=Miyakojima WG-20"  
XhoI 103 a 78 c 95 g 148 t  
BASE COUNT 103 a 78 c 95 g 148 t  
ORIGIN  
|||||  
alignment\_scores  
Quality: 582.00 Length: 141  
Ratio: 4.376 Gaps: 0  
Percent Similarity: 94.376 Percent Identity: 80.851

alignment\_block:  
US-09-271-584A-2 x AV408114 ..  
Align seg 1/1 to: AV408114 from: 1 to: 424

120 ILSerCysThrIleSerLeuGlyValThrGlnPhePheLysLysLeu 136  
1 ATATCTTGCCACCATATACTTTGGGTGCTAGCGAAGCTTTTAAGTGAAT 50  
136 vASpILeGlyThrPheAspLeuGlyAspTyrLeuAlaIlePhe 153  
51 AGATGTGGTCCACTGGAGTGGGATTTACTAGCAATGGTGAATAT 100  
153 heAlaLaThrAspSerValCysThrLeuGlnValLeuAsnGlnAspGlu 169  
101 TTGCTGTCACAGATTCTGTTGCACATTCAGGTGTTAAATCAAGATGAG 150  
170 ThrProLeuLeuTyrSerLeuValPheGlyGluGlyValValAsnAspAl 186  
151 ACACCTTTGCTGACAGCTTGTATTTGGGGAGGTGTGTGAATGATGC 200  
186 aThrSerValValPheAsnAlaIleGlnSerPheAspLeuThrHisL 203  
201 TATCATGCTGCTGCTTTTCAATCAATCAAAAGCTTTGACCTCAATCAAA 250  
203 euAsnHisGluAlaAlaPheHisLeuLeuGlyAsnPheLeuTyrLeuPhe 219  
251 TTAATCTTCAATGCTTTGCACTTTTGGGCAACTTCTGTATCTGTT 300  
220 LeuLeuSerThrLeuLeuGlyAlaAlaThrGlyLeuIleSerAlaTyrVa 236  
301 ATCGCAAGCAGCATGCTGGGGTTTGGACAGGTCTTCTCAGCTGATCAT 350  
236 lIleLysLysLeuTyrPheGlyArgHisSerThrAspArgGluAlaL 253  
351 CATTAAGAGCTATACATGGCAGGAGCTCTACAGATCGTGGTGGTCTC 400  
253 euMetMetLeuMetAlaTyrLeu 260  
|||||

401 TTATGATGCTAATGTCATACCTT 423  
seq\_name: gb\_est71:BE433982

seq\_documentation\_block:  
LOCUS BE433982 519 bp mRNA EST 24-JUL-2000  
DEFINITION EST7405060 tomato breaker fruit, TIGR Lycopersicon esculentum cDNA clone cLEG9M8, mRNA sequence.  
ACCESSION BE433982  
VERSION BE433982.1 GI:9431825  
KEYWORDS EST.  
SOURCE tomato.  
ORGANISM Lycopersicon esculentum  
Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Asteridae; euasterids 1; Solanales; Solanaceae; Solanum; Lycopersicon.

REFERENCE 1 (bases 1 to 519)  
AUTHORS Alcalá, J., Vrebalov, J., White, R., van der Hoeven, R.S., Holt, I.E., Liang, F., Hansen, T.S., Craven, M.B., Bowman, C.L., Renning, C.M., Niernman, W., Fraser, C.M., Martin, G.B., Giovannoni, J.J. and Tanksley, S.D.  
TITLE Generation of ESTs from tomato fruit tissue, breaker stage  
JOURNAL Unpublished (2000)  
COMMENT Contact: David Frisch  
Clemson University Genomics Institute  
Clemson University  
100 Jordan Hall, Clemson, SC 29634, USA  
Tel: 864 656 4366  
Fax: 864 656 4293  
Email: dfrisch@CLEMSON.EDU

FEATURES  
source  
1..519  
/organism="Lycopersicon esculentum"  
/cultivar="TA496"  
/db\_xref="taxon:4081"  
/clone="cLEG9M8"  
/clone\_lib="tomato breaker fruit, TIGR"  
/tissue\_type="pericarp"  
/dev\_stage="breaker"  
/lab\_host="SOLR"  
/note="vector: pBluescriptSKmCuadapt; Site\_1: EcoRI; Site\_2: XhoI; Fruit were harvested at the breaker stage (first sign of lycopene accumulation on the blossom end of the fruit). Fruit were cut in half and the seeds and locules were discarded prior to freezing the pericarp."  
XhoI 147 a 127 c 111 g 133 t 1 others  
BASE COUNT 147 a 127 c 111 g 133 t 1 others  
ORIGIN  
|||||

alignment\_scores  
Quality: 554.50 Length: 176  
Ratio: 3.798 Gaps: 3  
Percent Similarity: 82.955 Percent Identity: 62.500

alignment\_block:  
US-09-271-584A-2 x BE433982 ..  
Align seg 1/1 to: BE433982 from: 1 to: 519

359 ProLeuSerPheLeuSerAsnLeuAlaLysLysAsnGlnSerGluLysI 375  
1 CCATTATCACTATATCCAACTTAATGAAGAAATCTCGGAAGAAAAAT 50  
375 eAsnPheAsnMetClnValIleIleTyrPheSerGlyLeuMetArgGlyA 392  
51 TACCTTTAGGCAGCAACTGATATATGTTGGGAGGCTTTGATGAGAGCG 100  
392 laValSerMetAlaLeuAlaTyrAsnLysPheThrArgAlaGlyHisThr 408  
101 CAGTGTCCATGGCACTGGCATATAATAAGTTCACTCATGGGGGACACAT 150  
409 AspValArgGlyAsnAlaIleMetIleThrSerThrIleThrValCysLe 425  
|||||

```

151 CACTGCAGACATGCAATGATGATTACAGCAGCATACCACTGTTCT 200
|||||
425 uPheSerThrValValPheGlyMetLeuThrTyrsProLeuIleSerTyrL 442
|||||
201 ATTCAGCACAAATGGTATTTGGTCTTAATGACAAAACCCCTTATAAGTCTCC 250
|||||
442 euLeuProHisGlnAsnAlaThrThrsMetLeuSerAspAsnThr 458
|||||
251 TGCTGCCACACAGAGAGCAATGAGTACAGTGCATCAGGTGCAATAC 300
|||||
459 ProLysSerIleHisIleProLeuLeuAsp.....GlnAspSerPheII 473
|||||
301 CCTAAGTCTCTAACAGCCCACTCTAGGAGTCGAGAGGACTCT..... 345
|||||
473 eGluProSerGlyAsnHisAsnVal.....ProArgProAspSerI 487
|||||
346 .....GAAGGTGATTAAATGTCACAGATCTTCCTCAACCACCAAGTT 388
|||||
487 leArgGlyPheLeuThrArgProThrArgThrValHisTyrTyrTrpArg 503
|||||
389 TGAGGATGCTACTTACCGCACCAAGTCAATAAGTGCATCGGTACTGCGC 438
|||||
504 GlnPheAspSerPheMetArgProValPheGlyArgGlyPheVa 520
|||||
439 AAGTTTGACGATGATTCATCGCGCCCTATGTTGGTGTGCGGGATTGC 488
|||||
520 lProPheValProGlySerProThrGlu 529
|||||
489 TCCTNCTGCTCCTCGTCTCCACGGAA 516
|||||

```

seq\_name: gb\_est73:BE607262

```

seq_documentation_block: 401 bp mRNA EST 22-AUG-2000
LOCUS BE607262
DEFINITION NXCI_034_G06_F NXCI (Nsf Xylem Compression wood Inclined) Pinus
taeda cDNA clone NXCI_034_G06 5', mRNA sequence.
ACCESSION BE607262
VERSION BE607262.1 GI:9884342
KEYWORDS EST.
SOURCE loblolly pine.
ORGANISM Pinus taeda
Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
Coniferopsida; Coniferales; Pinaceae; Pinus; Pinus.
REFERENCE 1 (bases 1 to 401)
Soderoff,R.
Molecular Basis of Wood Formation in the Pine Megagenome
Unpublished (2000)
Contact: Johnson, Arthur
North Carolina State University
Tel: 919 515 7800
Fax: 919 515 7801
Email: ajohnson@unity.ncsu.edu
Seq primer: T3.

```

```

FEATURES
source
1..401
Location/Qualifiers
/organism="Pinus taeda"
/strain="Coastal plain loblolly pine from North Carolina"
/db_xref="taxon:3352"
/clone="NXCI_034_G06"
/clone_lib="NXCI (Nsf Xylem Compression wood Inclined)"
/tissue_type="Xylem"
/cell_type="Compression"
/dev_stage="Juvenile"
/lab_host="X11-Blue"
/note="Vector: Bluescript SK; Site_1: Eco RI; Site_2: XhoI
; The library is from early (spring) wood, taken from
three six-year old trees (three different genotypes), in
the juvenile phase. These trees were induced to form
compression wood by bending to a 45 degree angle and tying
them to the ground. Differentiating xylem was harvested
from the bottoms of the inclined stems, and a mixture of
all three genotypes was used for the library. oligo-dT

```

primed cDNA was directionally cloned into the EcoRI-XhoI  
 Bluescript SK vector arms. NOTE: The sequences contain a  
 'cDNA adapter' between the EcoRI site and the start of the  
 EST. The adapter sequence is 'AATTGGCAGCAG'."

BASE COUNT 98 a 84 c 79 g 125 t 15 others  
 ORIGIN

alignment\_scores:  
 Quality: 544.00 Length: 133  
 Ratio: 4.496 Gaps: 0  
 Percent Similarity: 90.977 Percent Identity: 77.444

alignment\_block:  
 US-09-271-584A-2 x BE607262 ..

Align seg 1/1 to: BE607262 from: 1 to: 401

```

213 GlyAsnPheLeuTyrLeuPheLeuSerThrLeuLeuGlyAlaAlaLath 229
|||||
2 GGTAACTTTGTTTCATCTCTTCTGACAAGCACTGCATTTGGGATTGCAGC 51
|||||
229 rGlyLeuIleSerAlaTyrValIleIleLysLeuTyrPheGlyArgHiss 246
|||||
52 TGGGCTTTTAAAGTGCATCATCATCAAGCTTTGTACTTTGGAAGGCAC 101
|||||
246 erThrAspArgGluValAlaLeuMetMetLeuMetAlaTyrLeuSerTyr 262
|||||
102 CTACTGATGCTGAAGTAGCAATTATGATACTCATGCGGTATCTCTCATAC 151
|||||
263 MetLeuAlaGluLeuPheAspLeuSerGlyIleLeuThrValPhePhecy 279
|||||
152 ATGCTGCTGAGCTGTTTTCACCTGAGTGGGATCTTGACTNNTTTCTCTN 201
|||||
279 sGlyIleValMetSerHisTyrThrTrpHisAsnValThrGluSerSera 296
|||||
202 NGGCATTTNNATGTCCTACACATGTCGATATGCTGACTGAAGCTC 251
|||||
296 rgIleThrThrLysHisThrPheAlaThrLeuSerPheLeuAlaGluThr 312
|||||
252 NNATTACCAACCAAGCATGCAATTTGCAACAATGTCAATTTATCGCTGAAACA 301
|||||
313 PheIlePheLeuTyrValGlyMetAspAlaLeuAspIleAspLysTrpAr 329
|||||
302 TTTATCTTCTTTATGTGGGAATGGATGCTCTGGACATTTGAAAAATGGAA 351
|||||
329 gSerValSerAspThrProGlyThrSerIleAlaValSerSerIleLeu 345
|||||
352 ATGGTGAAAGAGAGCCCTNNCACATCTGTTNNCTCAGCAGTGTCTC 400
|||||

```

seq\_name: gb\_est80:C91832

```

seq_documentation_block: 378 bp mRNA EST 20-APR-1998
LOCUS C91832
DEFINITION E31686_1A, mRNA sequence.
E31686_1A, mRNA sequence.
ACCESSION C91832
VERSION C91832.1 GI:3061200
KEYWORDS EST.
SOURCE Oryza sativa.
ORGANISM Oryza sativa.
Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
Magnoliophyta; Liliopsida; Poales; Poaceae; Oryza.
REFERENCE 1 (bases 1 to 378)
AUTHORS Sasaki,T. and Yamamoto,K.
TITLE Rice cDNA from panicle
JOURNAL Unpublished (1997)
COMMENT Contact: Takuji Sasaki
National Institute of Agrobiological Resources
Rice Genome Research Program
2-1-2 Kannondai,Tsukuba
Ibaraki,
Japan 305

```

Tel: 0298-38-7441

Fax: 0298-38-7468

Email: tsasakiaabr.affrc.go.jp.

Location/Qualifiers

## FEATURES

source

1..378

/organism="Oryza sativa"

/strain="Nipponbare"

/db\_xref="taxon:4530"

/clone="E31686\_1A"

/dev\_stage="rice panicle shorter than 3cm"

/note="Organ: panicle"

BASE COUNT 88 a 74 c 121 t 6 others

ORIGIN

## alignment\_scores:

Quality: 522.00 Length: 125

Ratio: 4.424 Gaps: 2

Percent Similarity: 94.400 Percent Identity: 83.200

## alignment\_block:

US-09-271-584A-2 x C91832

Align seg 1/1 to: C91832 from: 1 to: 378

238 LysLysLeuTyrPheGlyArgHisSerThrAspArgGluValAlaLeuMe 254

2 AAGAGCTATACATTTGGAGGCACTTCTACTGACCGTGAGGTGCCCTTAT 51

254 tMetLeuMetAlaTyrLeuSerTyrMetLeuAlaGluLeuPheAspLeuS 271

52 GATGCTCATGCTTACCTTTCATATATGCTGGCTGAGTTGCTAGATTGA 101

271 erGlyIleLeuThrValPhePheCysGlyIleValMetSerHisTyrThr 287

102 GCGGCACTTCTCACCCTATCTCTGTGTAATGTAATGTCACATTACACT 151

288 TrpHisAsnValThrGluSerSerArgIleThrThrLysHisThrPheAl 304

152 TGGCATACAGTTCACAGAGAGTTCAGAGTTACACAAAGCAGCATTTGC 201

304 aThrLeuSerPheLeuAlaGluThrPheIlePheLeuTyrValGlyMetA 321

202 AACTCTGCTCTTCAATGCTGAGACTTTTCTCTCTCTCTCTCTCTCTCT 251

321 spAlaLeuAspIleAspLysTrpArgSerValSerAspThrProGlyThr 337

252 ATGCATTGGATATTGAAAATGGGAGNTTNCAGTGACAGACCTGGNAAA 301

338 SerIleAlaValSerIleLeuMet.GlyLeuVal.MetValGlyArg 353

302 TCCATTNGGGTAAGCTCAATTTGCTAGGATGGTTCCCTGATTGGAGN 351

354 AlaAlaPheValPheProLeu 360

352 GCTGCTTTTGGNAATTCCGCTG 372

seq\_name: gb\_est48:AW694621

## seq\_documentation\_block:

LOCUS AW694621 757 bp mRNA EST 15-JUN-2000

DEFINITION NF078C11ST1F1085 Developing stem Medicago truncatula cDNA clone

NC\_078C11ST 5', mRNA sequence.

ACCESSION AW694621

VERSION AW694621.1 GI:7569383

KEYWORDS EST.

SOURCE barrel medic.

ORGANISM Medicago truncatula

Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;

Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I;

Fabales; Fabaceae; Papilionoideae; Medicago.

REFERENCE

1 (bases 1 to 757)

## AUTHORS

He,X.-Z., Shadle,G., Scott,A.D., Harris,A.R., Gonzales,R.A., Bell  
C.J., Flores,H.R., Inman,J.T., Weller,J.W., May,G.D. and Dixon  
R.A.

## TITLE

Expressed Sequence Tags from the Samuel Roberts Noble Foundation  
Medicago truncatula stem library

## JOURNAL

Unpublished (2000)

## COMMENT

Contact: Dixon RA

Plant Biology Division

The Samuel Roberts Noble Foundation

2510 Sam Noble Parkway, Ardmore, OK 73402, USA

Tel: 580 221 7302

Fax: 580 221 7380

Email: radixon@noble.org

Insert Length: 757 Std Error: 0.00

Plate: 078 row: C column: 11

Seq primer: TCACACAGGAAACAGCTATGAC.

## FEATURES

source

1..757

/organism="Medicago truncatula"

/db\_xref="taxon:3880"

/clone="NF078C11ST"

/clone\_lib="Developing stem"

/tissue\_type="stem"

/dev\_stage="Pooled developmental"

/note="Vector: Lambda Zap; Contains a mixture of

internodal stem segments"

BASE COUNT 230 a 171 c 143 g 213 t

## ORIGIN

## alignment\_scores:

Quality: 496.50 Length: 167

Ratio: 3.598 Gaps: 4

Percent Similarity: 82.635 Percent Identity: 61.677

## alignment\_block:

US-09-271-584A-2 x AW694621

Align seg 1/1 to: AW694621 from: 1 to: 757

334 ThrProGlyThrSerIleAlaValSerSerIleLeuMetGlyLeuValMe 350

2 AGTCCTGGAAACATCTATAGCAGCAAGTTCAGTATTTGTTGGGTCTAATACT 51

350 tValGlyArgAlaAlaPheValPheProLeuSerPheLeuSerAsnLeuA 367

52 TCTTGAAGAGCAGCGCTTTGTTTCCCTTATCCTTCTTATCCAACTTGA 101

367 laLysLysAsnGlnSerGluLysIleAsnPheAsnMetGlnValValIle 383

102 CTAAAAAATCACAACATCAGAACATCTCCTTCAGACAGCAAGTGATCAAT 151

384 TrpTrpSerGlyLeuMetArgGlyAlaValSerMetAlaLeuAlaTyrAs 400

152 TGGTGGCTGGTCTTATGAGAGTGCTGTTTCAATGGCACTTGGCTATAA 201

400 nLysPheThrArgAlaGlyHisThrAspValArgGlyAsnAlaIleMetI 417

202 TCAGTTCCACCATGTCGGGGCATACTCAACTGCTAGCAATGCAATCATGA 251

417 leThrSerThrIleThrValCysLeuPheSerThrValValPheGlyMet 433

252 TAACACAGCACCATCACTGTTGTCCTTTTCAGCACAGTGGTGGTTGGTTG 301

434 LeuThrLysProLeuIleSerTyrLeuLeuProHisGlnAsnAlaThrTh 450

302 CTGACTAAGCCACTTATAAGGCTTCTACTACCTCATCTCTCTCTCTCTCT 351

450 rSerMet....LeuSerAspAsnThrProLysSerIleHis..... 463

352 CAGCATGACACACACAGAGATCTACTACTCCAAAATACATTCATTCCTCCA 401

464 .....IleProLeuLeuAspGlnAsp.....SerPheIle 473



488 ATAACAATTATTTGTTGTTGCTGTTGGAACCTCTGATTTCCCTTGTAAAT 537



Percent Similarity: 94.690 Percent Identity: 76.106

alignment\_block:

US-09-271-584A-2 x AI941126

Align seg 1/1 to: AI941126 from: 1 to: 352

```
114 GlyAlaValGlyThrIleIleSerCysThrIleIleSerLeuGlyValTh 130
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
5 GGTGCTATTGCCACATTAATATCTGGTACCATCATACTTTGGGTGCCAC 54
130 rGlnPhePheLysLysLeuAspIleGlyThrPheAspLeuGlyAspTyrL 147
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
55 ACAAAATTTTAAAGAGGTTGGATGTTGGTCCACTGGCAATTAGGGGATTCC 104
147 euAlaIleGlyAlaIlePheAlaIleThrAspSerValCysThrLeuGln 163
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
105 TAGCATTTGGTGCAATATTGCTGCACGGATCTGTTGCACATTGCAG 154
164 ValLeuAsnGlnAspGluThrProLeuLeuTyrSerLeuValPheGlyG1 180
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
155 GTGCTTAATCAGGATGAGACACCTTTGCTGTACAGTCTTGTATTGGGA 204
180 uGlyValValAsnAspAlaThrSerValValPheAsnAlaIleGlnS 197
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
205 GGGTGTGTGAATGATGCTACATCAGTGGTCTTTTCAATGCAATCCAGA 254
197 erPheAspLeuThrHisLeuAsnHisGluAlaAlaPheHisLeuLeuGly 213
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
255 GCTTGACCTCAACCAATCGACCCCTCAATCGCTGGGCACCTTTTGGGA 304
214 AsnPheLeuTyrLeuPheLeuLeuSerThrLeuLeuGly 226
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
305 AATTCTGTATCTATTATTATGCAAGCACAAATGCTTGGGA 343
```

OM of: US-09-271-584A-2 to: N\_Geneseq\_36:\* out\_format : pfs  
Date: Mar 26, 2001 10:38 AM  
About: Results were produced by the GenCore software, version 4.5,  
Copyright (c) 1993-2000 Compugen Ltd.

Command line parameters:  
-MODEL=frameat.p2n.model -DEV=xlp  
-O=/cgn2\_1/USPTO.spool/US09271584/runat\_15032001\_151302.5706/app\_query.fasta\_1.604  
-DB=N\_Geneseq\_36 -QWTF=fastcap -SUFFIX=p2n.rng -GAPOP=12.000  
-GAPEXT=4.000 -MINMATCH=0.100 -LOOPCL=0.000 -LOOPEXT=0.000  
-FGAPOP=4.500 -FGAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500  
-DELOP=6.000 -DELEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500  
-TRANS=human40.cdi -LISP=45 -DOCALIGN=200 -THR\_SCORE=pct  
-THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pfs  
-NORM=ext -MINLEN=0 -MAXLEN=2000000000  
-USBR=US09271584.ecgn1\_1\_142 -NCPU=6 -ICPU=3 -LONGLOG -NO\_XLPXY  
-WAIT -THREADS=1

Search information block:

Query: US-09-271-584A-2  
Query length: 538  
Database: N\_Geneseq\_36:\*  
Database sequences: 480022  
Database length: 187831343  
Search time (sec): 93.960000

| score_list: | Sequence   | Strd Orig | zScore   | EScore  | Len      | Documentation |
|-------------|--|-----------|----------|---------|----------|---------------|
|             | /SID56/gcgdata/geneseq/geneseq/NA1999.DAT:222591 + |           | 2.6e-279 | 2178    | 1 A      |               |
|             | /SID56/gcgdata/geneseq/geneseq/NA1999.DAT:222595 + |           | 1.3e-275 | 2284    | 1 A      |               |
|             | /SID56/gcgdata/geneseq/geneseq/NA1999.DAT:222592 + |           | 1570.00  | 2924.52 | 4.1e-155 | 1788          |
|             | /SID56/gcgdata/geneseq/geneseq/NA1999.DAT:222606 + |           | 522.00   | 968.46  | 3.7e-46  | 378           |
|             | /SID56/gcgdata/geneseq/geneseq/NA1999.DAT:222602 + |           | 512.50   | 935.33  | 3.4e-44  | 1683          |
|             | /SID56/gcgdata/geneseq/geneseq/NA1999.DAT:222630 + |           | 463.00   | 835.15  | 9.9e-39  | 2584          |
|             | /SID56/gcgdata/geneseq/geneseq/NA1999.DAT:222611 + |           | 441.00   | 788.24  | 4.1e-36  | 4151          |
|             | /SID56/gcgdata/geneseq/geneseq/NA1999.DAT:222603 + |           | 415.00   | 768.48  | 5.1e-35  | 330           |
|             | /SID56/gcgdata/geneseq/geneseq/NA1999.DAT:222593 + |           | 401.00   | 720.64  | 2.4e-32  | 2122          |
|             | /SID56/gcgdata/geneseq/geneseq/NA1999.DAT:222597 + |           | 247.00   | 443.13  | 6.8e-17  | 714           |
|             | /SID56/gcgdata/geneseq/geneseq/NA1999.DAT:222598 + |           | 246.00   | 419.74  | 1.4e-15  | 4605          |
|             | /SID56/gcgdata/geneseq/geneseq/NA1999.DAT:222607 + |           | 201.50   | 329.07  | 1.5e-10  | 8339          |
|             | /SID56/gcgdata/geneseq/geneseq/NA1999.DAT:222602 + |           | 190.00   | 347.07  | 1.5e-11  | 268           |
|             | /SID56/gcgdata/geneseq/geneseq/NA1999.DAT:222607 + |           | 185.50   | 291.89  | 1.8e-08  | 15363         |
|             | /SID56/gcgdata/geneseq/geneseq/NA1999.DAT:222607 + |           | 166.50   | 284.64  | 4.6e-08  | 1294          |
|             | /SID56/gcgdata/geneseq/geneseq/NA1999.DAT:222607 + |           | 144.00   | 228.16  | 6.4e-05  | 4392          |
|             | /SID56/gcgdata/geneseq/geneseq/NA1999.DAT:222607 + |           | 144.00   | 228.16  | 6.4e-05  | 4392          |
|             | /SID56/gcgdata/geneseq/geneseq/NA1999.DAT:222607 + |           | 143.50   | 240.86  | 1.3e-05  | 1347          |
|             | /SID56/gcgdata/geneseq/geneseq/NA1999.DAT:222607 + |           | 142.00   | 155.89  | 0.6787   | 1664976       |
|             | /SID56/gcgdata/geneseq/geneseq/NA1999.DAT:222607 + |           | 137.00   | 233.30  | 3.3e-05  | 1244          |
|             | /SID56/gcgdata/geneseq/geneseq/NA1999.DAT:222607 + |           | 137.00   | 225.86  | 8.6e-05  | 1710          |
|             | /SID56/gcgdata/geneseq/geneseq/NA1999.DAT:222607 + |           | 135.50   | 242.19  | 1.1e-05  | 325           |
|             | /SID56/gcgdata/geneseq/geneseq/NA1999.DAT:222607 + |           | 132.00   | 197.74  | 0.0032   | 8654          |
|             | /SID56/gcgdata/geneseq/geneseq/NA1999.DAT:222607 + |           | 129.00   | 214.51  | 0.0004   | 1239          |
|             | /SID56/gcgdata/geneseq/geneseq/NA1999.DAT:222607 + |           | 129.00   | 191.22  | 0.0073   | 9326          |
|             | /SID56/gcgdata/geneseq/geneseq/NA1999.DAT:222607 + |           | 122.00   | 178.90  | 0.0355   | 8654          |
|             | /SID56/gcgdata/geneseq/geneseq/NA1999.DAT:222607 + |           | 118.00   | 160.56  | 0.3729   | 22080         |
|             | /SID56/gcgdata/geneseq/geneseq/NA1999.DAT:222607 + |           | 117.50   | 185.53  | 0.0152   | 2337          |
|             | /SID56/gcgdata/geneseq/geneseq/NA1999.DAT:222607 + |           | 117.50   | 173.23  | 0.0734   | 6785          |
|             | /SID56/gcgdata/geneseq/geneseq/NA1999.DAT:222607 + |           | 116.00   | 188.54  | 0.0103   | 1409          |
|             | /SID56/gcgdata/geneseq/geneseq/NA1999.DAT:222607 + |           | 113.00   | 184.51  | 0.0173   | 1224          |
|             | /SID56/gcgdata/geneseq/geneseq/NA1999.DAT:222607 + |           | 113.00   | 106.71  | 372.34   | 1038602       |
|             | /SID56/gcgdata/geneseq/geneseq/NA1999.DAT:222607 + |           | 111.50   | 159.45  | 0.4298   | 8411          |
|             | /SID56/gcgdata/geneseq/geneseq/NA1999.DAT:222607 + |           | 111.00   | 179.62  | 0.0324   | 1350          |
|             | /SID56/gcgdata/geneseq/geneseq/NA1999.DAT:222607 + |           | 109.50   | 176.44  | 0.0487   | 1392          |
|             | /SID56/gcgdata/geneseq/geneseq/NA1999.DAT:222607 + |           | 109.50   | 176.44  | 0.0487   | 1392          |
|             | /SID56/gcgdata/geneseq/geneseq/NA1999.DAT:222607 + |           | 109.50   | 175.21  | 0.0570   | 1427          |
|             | /SID56/gcgdata/geneseq/geneseq/NA1999.DAT:222607 + |           | 108.50   | 153.18  | 0.9607   | 8876          |
|             | /SID56/gcgdata/geneseq/geneseq/NA1999.DAT:222607 + |           | 108.50   | 153.08  | 0.9729   | 8952          |
|             | /SID56/gcgdata/geneseq/geneseq/NA1999.DAT:222607 + |           | 107.50   | 150.26  | 1.40     | 9711          |

/SID56/gcgdata/geneseq/geneseq/NA1999.DAT:222591 + 89.82 3.3e+03 183012  
/SID56/gcgdata/geneseq/geneseq/NA1999.DAT:222591 + 106.50 159.82 0.4099 3601  
/SID56/gcgdata/geneseq/geneseq/NA1999.DAT:222591 + 105.50 161.58 0.3273 2627  
/SID56/gcgdata/geneseq/geneseq/NA1999.DAT:222591 + 104.50 143.60 3.28 10594  
/SID56/gcgdata/geneseq/geneseq/NA1999.DAT:222591 + 104.50 142.86 3.61 11303

seq\_name: /SID56/gcgdata/geneseq/geneseq/NA1999.DAT:222591

seq\_documentation\_block:

ID Z22591 standard; cDNA; 2178 BP.  
XX Z22591;  
XX 18-JAN-2000 (first entry)  
DT Arabidopsis thaliana Na/H transporter gene ALNHX1.  
DE Sodium; proton; antiport; transporter; salt tolerance; salt management;  
KW transgenic plant; survival; soil; farming; accumulation; irrigation;  
KW crop; ss.  
XX Arabidopsis thaliana.  
XX WO9947679-A2.  
XX 23-SEP-1999.  
XX 18-MAR-1999; 99WO-CA00219.  
XX 18-MAR-1998; 98US-0078474.  
XX 15-JAN-1999; 99US-0116111.  
XX (BLUM/) BLUMWALD E.  
XX (APSE/) APSE M.  
XX (SNED/) SNEDDEN W.  
XX (AHAR/) AHARON G.  
XX Blumwald E, Apse M, Snedden W, Aharon G;  
WPI; 1999-571840/48.  
P-PSDB; Y40901.  
XX Nucleic acid molecules encoding sodium/proton transport polypeptides,  
useful in genetic engineering salt tolerance in crop plants -  
Claim 4; Fig 1A; 93pp; English.

The invention relates to an isolated nucleic acid molecule encoding a plant Na/H antiport (PNHX) transporter polypeptide, or a fragment and capable of increasing salt tolerance in a cell. This sequence corresponds to the gene encoding the ALNHX1 transporter from Arabidopsis thaliana. The Na/H transporter polypeptides provide a means of intracellular salt management, particularly in plants. The sequences are useful for producing transgenic plants that are capable of surviving in soil with high salt levels that would normally inhibit growth of the crop species. This would be useful in farming land in areas that are generally considered unproductive through salt accumulation and poor irrigation, e.g. in India, Australia, and prairies in USA or Canada. Commercial crops, such as potatoes, tomatoes, brassica, cotton, sunflower, strawberries, spinach, lettuce, rice, soybean, corn, wheat, barley, alfalfa, sorghum, alfalfa, salicornia and others would benefit from increased salt tolerance.

Sequence 2178 BP; 576 A; 413 C; 472 G; 717 T; 0 other;

alignment\_scores:

Quality: 2755.00 Length: 538  
Ratio: 5.121 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000  
alignment\_block:  
US-09-271-584A-2 x Z22591

Align seg 1/1 to: Z22591 from: 1 to: 2178

```
1 MetLeuAspSerLeuValSerLysLeuProSerLeuSerThrSerAspHi 17
286 ATGTTGGATTCTCTAGTGTGCGAACCTGCCCTCGTTATCGACATCTGAICA 335
17 sAlaSerValValAlaLeuAsnLeuPheValAlaLeuLeuCysAlaCysI 34
336 CGCTTCGTGGTTCGGTTGAATCTCTTTGTTGGACATCTTTTGTGCTTGTGA 385
34 leValLeuGlyHisLeuLeuGluGlnAsnArgTrpMetAsnGluSerIle 50
386 TTGTTCTTGGTCATCTTTTGAAGAGAATAGATGATGAACGAATCCATC 435
51 ThrAlaLeuLeuLeuGlyLeuGlyThrGlyValThrIleLeuLeuIleSe 67
436 ACCCCCTTGTGATTTGGGCTAGGCATGCTGGTTACCAATTTTGTGTGATTAG 485
67 rLysGlyLysSerSerHisLeuLeuValPheSerGluAspLeuPheIle 84
486 TAAAGGAAAGAGCTCGCATCTTCTCGTCTTTAGTGAAGATCTTTTCTTCA 535
84 leftyLeuLeuProProIleIlePheAsnAlaGlyPheGlnValLysLys 100
536 TATATCTTTTCCCAACCATATATTCATCAATGACAGGTTTCAAGTAAAAAG 585
101 LysGlnPhePheArgAsnPheValThrIleMetLeuPheGlyAlaValGI 117
586 AAGCAGATTTTCCCAATTTCCGTCGACATATATGCTTTTGGTGTGCTGGTGG 635
117 yThrIleLeuSerCysThrIleLeuSerLeuGlyValThrGlnPhePheL 134
636 GACTATTATTTTGGCAACATCATATCTTAGGTTGTAACAGATCTTTTA 685
134 ysLysLeuAspIleGlyThrPheAspLeuGlyAspTyrLeuAlaIleGly 150
686 AGAAGTTGGACATTTGGAAACCTTTGACTTGGGTGATTATCTTGCTATTGGT 735
151 AlaIlePheAlaIaThrAspSerValCysThrLeuGlnValLeuAsnGI 167
736 GCCATATTTGTCACACAGATTACATCTAGTGTGACACTGCAGGTCTCGAATCA 785
167 nAspGluThrProLeuLeuFyrSerLeuValPheGlyGluGlyValVala 184
786 AGACGAGACACCTTTGCTTTACAGCTCTGTATTTCGAGAGGGTGTGTGA 835
184 snAspAlaThrSerValValPheAsnAlaIleGlnSerPheAspLeu 200
836 ATGATGCAACGTCAGTTGTGCTCTTCAACGCGATTCAGAGCTTTGTATCTC 885
201 ThrHisLeuAsnHisGluAlaAlaPheHisLeuLeuGlyAsnPheLeuTy 217
886 ACTCACCTTAACACGAGAGCTGCTTTCACTCTTGGAACTCTTGTGA 935
217 rLeuPheLeuLeuSerThrLeuLeuGlyAlaAlaThrGlyLeuIleSerA 234
936 TTTGTTTCTCCTAAGTACCTTGTGCTGGTGTGCTGCAACCGGCTGTGATAAGTG 985
234 laTyrValIleIleLysLysLeuTyrPheGlyArgHisSerThrAspArgGlu 250
986 CGTATGTTATCAAGAAGCTATACTTTGGAAGGCACCTCACTGACCGAGAG 1035
251 ValAlaLeuMetMetLeuMetAlaTyrLeuSerTyrMetLeuAlaGluLe 267
1036 GTTGCCCTTATGATGCTTATGCGGTACTTCTTATATGCTTGTGAGCT 1085
267 upheAspLeuSerGlyIleLeuThrValPhePheCysGlyIleValMetS 284
1086 TTTTCGACTTGACGGGTATCCCTCACTGTGTTTTTCTGCTGATTTGTGATGT 1135
284 erHisTyrThrTrpHisAsnValThrGluSerSerArgIleThrThriLys 300
|||||
```

```
1136 CCCATTACACATGGCACAAATGTAAACGGAGAGCTCAAGAAATAACACAAAG 1185
301 HisThrPheAlaThrLeuSerPheLeuAlaGluThrPheIlePheLeuTy 317
|||||
1186 CATACCTTTTGAACCTTTGTCATTTCTTCGGGAGACATTTATTTCTTGTA 1235
317 rValGlyMetAspAlaLeuAspIleAspLysTrpArgSerValSerAspT 334
|||||
1236 TGTGTGAATGGATGCTTGGCATTGACAAGTGGAGATCCGTGAGTGACA 1285
334 hrProGlyThrSerIleAlaValSerSerIleLeuMetGlyLeuValMet 350
|||||
1286 CACCGGGAACATCGATCGAGTGCAGTCAATCCCTAATGGCTGGTCATG 1335
351 ValGlyArgAlaAlaPheValPheProLeuSerPheLeuSerAsnLeuAl 367
|||||
1336 GTTGAAGAGCAGCGTTCGTCTTTCGGTTATCGTTTCTATCTAATTAGC 1385
367 alysLysAsnGlnSerGluLysIleAsnPheAsnMetGlnValValIleT 384
|||||
1386 CAAGAAGAATCAAGCGAGAAATCAACTTTAAACATGCAAGTGTGTGATT 1435
384 rPTrpSerGlyLeuMetArgGlyAlaValSerMetAlaLeuAlaTyrAsn 400
|||||
1436 GGTGTGCTGTCTCATGAGAGGTGCTGTATCTATGCGCTCTTGATACAC 1485
401 LysPheThrArgAlaGlyHisThrAspValArgGlyAsnAlaIleMetI 417
|||||
1486 AAGTTTACAAGGCGCGGCACACAGATGTACGCGGAATGCAATCATGAT 1535
417 ehThrSerThrIleThrValCysLeuPheSerThrValValPheGlyMetL 434
1536 CACGAGTAGCATAACTGTCTGCTTTTGTAGCACAGTGTGTGTGTATGC 1585
434 euThrLysProLeuIleSerTyrLeuLeuProHisGlnAsnAlaThrThr 450
1586 TGACCAACCACTCATAGTACCTATTACCGCACCAACGCCACACAGG 1635
451 SerMetLeuSerAspAsnThrProLysSerIleHisIleProLeuLe 467
1636 AGCATGTTATCTGATGACAAACACCCCAAAATCCATACATATCCCTTGT 1685
467 uAspGlnAspSerPheIleGluProSerGlyAsnHisAsnValProArgP 484
1686 GGACCAAGACTCGTTCAATGAGCTTTCAGGAACCAACAAATGTCCTCGC 1735
484 roAspSerIleArgGlyPheLeuThrArgProThrArgThrValHisTyr 500
1736 CTGACAGTATACGTGCTTCTTGACACGGCCCACTCGAACCGTGCAATTAC 1785
501 TyrTrpArgGlnPheAspAspSerPheMetArgProValPheGlyGlyAr 517
1786 TACTGGAGACAAATTTGATGACTCTCTTCATGCGACCGCTCTTTTGGAGTGC 1835
517 gGlyPheValProPheValProGlySerProThrGluArgAsnProProA 534
1836 TGCGTTTGTACCCCTTTGTTCCAGGTTCTCCAACACTGAGAAACCCCTCCG 1885
534 spLeuSerLysAla 538
1886 ATCTTAGTAAGGCT 1899
seq_name: /SID56/gcgdata/geneseq/geneseq/NA1999.DAT:22595
seq_documentation_block:
ID Z22595 standard; cDNA; 2884 BP.
XX
AC Z22595;
XX
DT 18-JAN-2000 (first entry)
XX
DE Arabidopsis thaliana Na/H transporter gene.
XX
```

KW Sodium; proton; antiport; transporter; salt tolerance; salt management;  
 KW transgenic plant; survival; soil; farming; accumulation; irrigation;  
 KW crop; ss.

XX Arabidopsis thaliana.

XX WO9947679-A2.

XX 23-SEP-1999.

XX 18-MAR-1999; 99WO-CA00219.

XX 18-MAR-1998; 98US-0078474.

XX 15-JAN-1999; 99US-0116111.

XX (BLUM/) BLUMWALD E.

XX (APSE/) APSE M.

XX (SNED/) SNEDDEN W.

XX (AHAR/) AHARON G.

XX Blumwald E, Apse M, Snedden W, Aharon G;

XX WPI; 1999-571840/48.

XX P-PSDB; Y40905.

XX Nucleic acid molecules encoding sodium/proton transport polypeptides,  
 useful in genetic engineering salt tolerance in crop plants

XX Claim 48; Fig 5A-B; 93pp; English.

XX The invention relates to an isolated nucleic acid molecule encoding  
 a plant Na/H antiport (PNHX) transporter polypeptide, or a fragment  
 and capable of increasing salt tolerance in a cell. This sequence  
 corresponds to the gene encoding a transporter from Arabidopsis thaliana.  
 CC The Na/H transporter polypeptides provide a means of intracellular salt  
 management, particularly in plants. The sequences are useful for  
 CC producing transgenic plants that are capable of surviving in soil with  
 high salt levels that would normally inhibit growth of the crop species.  
 CC This would be useful in farming land in areas that are generally  
 considered unproductive through salt accumulation and poor irrigation,  
 CC e.g. in India, Australia, and prairies in USA or Canada. Commercial  
 CC crops, such as potatoes, tomatoes, brassica, cotton, sunflower,  
 CC strawberries, spinach, lettuce, rice, soybean, corn, wheat, rye, barley,  
 CC atriplex, sorghum, alfalfa, salicornia and others would benefit from  
 CC increased salt tolerance.

XX Sequence 2284 BP; 589 A; 444 C; 485 G; 766 T; 0 other;

#### alignment\_scores:

Quality: 2720.00 Length: 540  
 Ratio: 5.065 Gaps: 2  
 Percent Similarity: 99.444 Percent Identity: 99.444

#### alignment\_block:

US-09-271-584A-2 x 222595

Align seg 1/1 to: 222595 from: 1 to: 2284

1 MetLeuAspSerLeuValSerLysLeuProSerLeuSerThrSerAspHi 17  
 401 ATCTTGGATTCTCTAGTGTGCGAACTGCCCTTCGTTATCGACATCGATCA 450  
 17 sAlaSerValValAlaLeuAsnLeuPheValAlaLeuLeuCysAlaCysI 34  
 451 CGCTTCTGTGGTGGCTTGAATCTCTTTGTGCACCTTCTTTGTGCTTGA 500  
 34 leValLeuGlyHisLeuLeuGluGluAsnArgTrpMetAsnGluSerile 50  
 501 TTGTTCTTGGTCATCTTTTGGAGAGATAGATGGATGAACCAATCCATC 550  
 51 ThrAlaLeuLeuIleGlyLeuGlyThrGlyValThrIleLeuLeuIleSe 67  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||

551 ACCGCTTGTGATTTGGGCTAGGCACCTGGTCTTACCATTTTTGTGATTAG 600  
 67 rLysGlyLysSerSerHisLeuLeuValPheSerGluAspLeuPhePheI 84  
 601 TAAAGGAAAAAGCTCGCATCTTCTCGTCTTTAGTGAAGATCTTTTCTTCA 650  
 84 leTyrLeuLeuProProIleIlePheAsnAlaGlyPheGlnValLysLys 100  
 651 TATATCTTTTGGCACCATTATATCAATGACGGTTTCAAGTAAAAAAG 700  
 101 LysGlnPhePheArgAsnPheValThrIleMetLeuPheGlyAlaValGI 117  
 701 AAGCAGTTTTTCCGCAATTTGCTGACTATTATGCTTTTGGTGTGTTGG 750  
 117 yThrIleIleSerCysThrIleIleSerLeuGlyValThrGlnPhePheI 134  
 751 GACTATTATTCTTGCACAATCATATCTTAGTGTACACAGTCTTTTA 800  
 134 yLysLeuAspIleGlyThrPheAspLeuGlyAspTyrLeuAlaIleGly 150  
 801 AGAAGTTGGACATTGGAACTTTGACTTTGGGTGATTATCTTGTCTATTGGT 850  
 151 AlaIlePheAlaAlaThrAspSerValCysThrLeuGlnValLeuAsnGI 167  
 851 GCCATATTGCTGCAACAGATTCAGTATGACATGCAGGTTCTGANTCA 900  
 167 nAspGluThrProLeuLeuTyrSerLeuValPheGlyGluGlyValVala 184  
 901 AGACGAGACACCTTTGCTTTACAGTCTTGTATTGCGAGAGGGTGTGTGA 950  
 184 snAspAlaThrSerValValValPheAsnAlaIleGlnSerPheAspLeu 200  
 951 ATGATGCAACGTCAGTGTGGTCTTCAACGGGATTCAGAGCTTTGATCTC 1000  
 201 ThrHisLeuAsnHisGluAlaAlaPheHisLeuLeuGlyAsnPheLeuTy 217  
 1001 ACTCACTAAACACGAGAGTCTTTTTCATCTTCTTGGAAACTTTCTGTA 1050  
 217 rLeuPheLeuLeuSerThrLeuLeuGlyValAlaAlaThrGlyLeuIleSe 234  
 1051 TTGTGTTCTTAAAGTACCTTGTGGTGTGCAACCGCTGTGATAAGTG 1100  
 234 laTyrValIleLysLysLeuTyrPheGlyArgHisSerThrAspArgGlu 250  
 1101 CGTATGTTATCAAGAGCTATATCTTTGGAAGGCATCAACTGACCGAGAG 1150  
 251 ValAlaLeuMetMetLeuMetAlaTyrLeuSerTyrMetLeuAlaGluLe 267  
 1151 GTTGCCCTTATGCTTATGCGGTATCTTTCTTATATGCTTGTGAGCT 1200  
 267 upheAspLeuSerGlyIleLeuThrValPhePheCysGlyIleValMets 284  
 1201 TTTTCGACTTGAAGGGTATCTCTCCTCCTGTTTCTTGTGTTGTGATGT 1250  
 284 erHisTyrThrTrpHisAsnValThrGluSerSerArgIleThrThrLys 300  
 1251 CCATTACACATGGCAATGTAAAGGAGCTCAAGATACACAAAG 1300  
 301 HisThrPheAlaThrLeuSerPheLeuAlaGluThrPheIlePheLeuTy 317  
 1301 CATACCTTTTGGCAACTTTGTCAATTTCTTGGGAGACATTTATTTCTTGA 1350  
 317 rValGlyMetAspAlaLeuAspIleAspLysTrpArgSerValSerAspT 334  
 1351 TGTGGAATGGATGGCTTTGGACATTTGACAAGTGGAGATCCGTTGAGTGACA 1400  
 334 hrProGlyThrSerIleAlaValSerSerIleLeuMetGlyLeuValMet 350  
 1401 CACCGGGAACATCGATCGCAGTGAAGTCAATCTTAATGGTCTGTGTCATG 1450  
 351 ValGlyArgAlaAlaPheValPheProLeuSerPheLeuSerAsnLeuAl 367  
 1451 GTTGAAGAGACGCTGCTCTTTCCGTTATCGTTTCTATCTAATCTTAGC 1500





CC species. This would be useful in farming land in areas that are generally  
 CC considered unproductive through salt accumulation and poor irrigation,  
 CC e.g. in India, Australia, and prairies in USA or Canada. Commercial  
 CC crops, such as potatoes, tomatoes, brassica, cotton, sunflower,  
 CC strawberries, spinach, lettuce, rice, soybean, corn, wheat, rye, barley,  
 CC atriplex, sorghum, alfalfa, salicornia and others would benefit from  
 CC increased salt tolerance.  
 XX  
 SQ Sequence 378 BP; 88 A; 74 C; 89 G; 121 T; 6 other;

alignment\_scores:  
 Quality: 522.00 Length: 125  
 Ratio: 4.424 Gaps: 2  
 Percent Similarity: 94.400 Percent Identity: 83.200

alignment\_block:

US-09-271-584A-2 x 222606 ..

Align seg 1/1 to: 222606 from: 1 to: 378

238 LysLysLeuTyrPheGlyArgHisSerThrAspArgGluValAlaLeuMe 254  
 |||||  
 2 AAGAACTATACATTGGAAGGCATTCTACTGACCGTTCAGGTTGCCCTTAT 51  
 |||||  
 254 tMetLeuMetAlaTyrLeuSerTyrMetLeuAlaGluLeuPheAspLeuS 271  
 |||||  
 52 GATGCTCATCGCTTACCTTCATATATGCTGGCTGAGTGTGCTAGATTGA 101  
 |||||  
 271 erGlyLeuLeuThrValPhePheCysGlyLeValMetSerHisTyrThr 287  
 |||||  
 102 GCGCAATTCACCGTATCTCTGTTGGTATTGTAATGTCACATTACACT 151  
 |||||  
 288 TrpHisAsnValThrGluSerSerArgIleThrThrLysHisThrPheAl 304  
 |||||  
 152 TGCATTAACGTCACAGAGTTCAAGAGTTTACAAAGACGACCATTTGC 201  
 |||||  
 304 aThrLeuSerPheLeuAlaGluThrPheIlePheLeuTyrValGlyMetA 321  
 |||||  
 202 AACTCTGCTCTTCATGCTGAGACTTTCTCTCTGATGTTGGATGG 251  
 |||||  
 321 spAlaLeuAspIleAspLysTrpArgSerValSerAspThrProGlyThr 337  
 |||||  
 252 ATGCATTGGATATTGAAAAATGGGAGNTTNCACAGTACACCTGGNAA 301  
 |||||  
 338 SerIleAlaValSerSerIleLeuMet.GlyLeuValMetValGlyArg 353  
 |||||  
 302 TCCATTNGGTAAGCTCAATTTGCTAGGAGTTGGTTCCTGATTGGAAGN 351  
 |||||  
 354 AlaAlaPheValPheProLeu 360  
 |||||  
 352 GCTGCTTTTGNAAATCCCTTG 372

seq\_name: /SIDS6/gcgdata/geneseq/geneseqn/NA1999.DAT:222602

seq\_documentation\_block:

ID 222602 standard; cDNA; 1683 BP.

XX

AC 222602;

XX

DT 18-JAN-2000 (first entry)

XX

DE Arabidopsis thaliana Na/H transporter gene AtNHX4 variant #1.

XX

KW Sodium; proton; antiport; transporter; salt tolerance; salt management;  
 KW transgenic plant; survival; soil; farming; accumulation; irrigation;  
 KW crop; ss.

OS Arabidopsis thaliana.

XX

PN W09947679-A2.

XX

PD 23-SEP-1999.

XX 18-MAR-1999; 99WO-CA00219.  
 PF  
 XX 18-MAR-1998; 98US-0078474.  
 PR  
 XX 15-JAN-1999; 99US-0116111.  
 XX  
 PA (BLUM/) BLUMWALD E.  
 PA (APSE/) APSE M.  
 PA (SNED/) SNEDDEN W.  
 PA (AHAR/) AHARON G.  
 XX  
 XX Blumwald E, Apse M, Snedden W, Aharon G;  
 XX WPI; 1999-571840/48.  
 DR  
 DR P-PSDB; Y40906.  
 XX  
 XX Nucleic acid molecules encoding sodium/proton transport polypeptides,  
 PT useful in genetic engineering salt tolerance in crop plants -  
 XX  
 XX Claim 4; Fig 1D; 93pp; English.  
 XX  
 CC The invention relates to an isolated nucleic acid molecule encoding a  
 CC plant Na/H antiport (PNHX) transporter polypeptide, or a fragment and  
 CC capable of increasing salt tolerance in a cell. This sequence corresponds  
 CC to one of 2 splice variants of the AtNHX4 transporter gene from  
 CC Arabidopsis thaliana. The Na/H transporter polypeptides provide a means  
 CC of intracellular salt management, particularly in plants. The sequences  
 CC are useful for producing transgenic plants that are capable of surviving  
 CC in soil with high salt levels that would normally inhibit growth of the  
 CC crop species. This would be useful in farming land in areas that are  
 CC generally considered unproductive through salt accumulation and poor  
 CC irrigation, e.g. in India, Australia, and prairies in USA or Canada.  
 CC Commercial crops, such as potatoes, tomatoes, brassica, cotton,  
 CC sunflower, strawberries, spinach, lettuce, rice, soybean, corn, wheat,  
 CC rye, barley, atriplex, sorghum, alfalfa, salicornia and others would  
 CC benefit from increased salt tolerance.  
 XX  
 SQ Sequence 1683 BP; 392 A; 360 C; 367 G; 564 T; 0 other;  
 alignment\_scores:  
 Quality: 512.50 Length: 536  
 Ratio: 1.567 Gaps: 21  
 Percent Similarity: 61.007 Percent Identity: 30.410  
 alignment\_block:  
 US-09-271-584A-2 x 222602 ..  
 Align seg 1/1 to: 222602 from: 1 to: 1683  
 18 AlaSerValValAlaLeuAsnLeuPheValAlaLeuLeuCysAlaCysI 34  
 |||||  
 37 GCGGCGCGGTGGTACTGCTTCAGATTATGATGCTGCTGCTTCCTT 86  
 |||||  
 34 eValLeuGlyHisLeuLeuGluGluAsnArg.....TrpMetAsnGluS 49  
 |||||  
 87 GGTCTCTGCGCATGTCTCCGCGCTCATCGATTCCACTATCTCTCTGAAG 136  
 |||||  
 49 erIleThrAlaLeuLeuIleGlyLeuGlyThrGlyValThrIleLeuLeu 65  
 :: |||||  
 137 CCAGCGGTTCGCTCTCATTTGTTTAATCGTCGGTATA...CTTGCTAAT 183  
 |||||  
 66 IleSerLysGlyLysSerSer.....HisLeuLeuValPheSerGluAs 80  
 |||||  
 184 ATCTCCGATAGTACAGATTAGGACGTGTTTAAATTTCCACGAAGA 233  
 |||||  
 80 pLeuPheHeiLeuLeuLeuProIleIlePheAsnAlaGlyPheG 97  
 ::|||  
 234 GTCTCTCTCTGTTTGTTCCTCCCATCATATTCAGTCAGGTTTCA 283  
 |||||  
 97 InValLysLysLysGlnPhePheArgAsnPheValThrIleMetLeuPhe 113  
 ::|||  
 284 GTCTTCAACCTAAACCATTCCTTTTAACTTTGGAGCCATTGTTACCTTT 333



397 LeuAla.TyrAsnLysPheThrArGAlaGlyHisThrAspValArgGlyA 413  
 1180 CTTGCATTCATCAATCAGTCTGATCTACGAGGGTCACGGCCAAATCAT 1229  
 413 snAlaIleMetIleThrSerThrIleThrValCysLeuPheSerThrVal 429  
 1230 CT.....TTACTGCAAAACCACATATTGTTGTCACGGTTTTTA 1270  
 430 ValPheGlyMetLeuThrLysProLeuIleSerTyrLeuLeuProHisGI 446  
 1271 CTAATAGAGGTCGACAGGTAAATGTTGGAAGCTTTG..... 1309  
 446 nAsnAlaThrThrSerMetLeuSerAspAsnThrProLysSerIleH 463  
 1310 .....GAAGTTGTAGGTGACGAT..... 1327  
 463 isIleProLeuLeuAspGlnAspSerPhe.....IleGluProSer 476  
 1328 .....CTTGAT...GACTCCATGTCTGAAGGCTTTGAAGAGAC 1363  
 477 GlyAsnHisAsnValProArgProAspSerIleArgGlyPheLeuThrAr 493  
 1364 GATCATCAGTATGTCCTCTCTTTAGCATGGAGCTTCATCTGACGA 1413  
 493 gProThr.....ArgThrValH 499  
 1414 GGATACATCATCATCAGAGCAGGTTCAAGATCAAGCTCAAGGAGGTTTC 1463  
 499 isTyrTyrArgGlnPhe.....AspAspSerPheMetArgPro 512  
 1464 ACAAAACCATACATCATTCACCGCGTTGGACAAAAACTTTCTGACTCCG 1513  
 513 ValPhe 514  
 1514 TTCTTC 1519

seq\_documentation\_block:

XX AC T51330;

Human Na<sup>+</sup>/H<sup>+</sup> exchanger isoform NHE3 composite cDNA.

OS Homo sapiens.

|    |     |          |
|----|-----|----------|
| FT | CDS | 11..2515 |
| FT | CDS | 11..2515 |

FT  
/note= "codon CAC (bases 1870-1972) cod

XX

PF 10-JUL-1996; 96WO-US11670.

XX

PI Brant SR, Donowitz M, Tse C, Yun CHC;

DR WPI; 1997-119051/11.





[illegible]

seq name: /SIDSS6/gcadata/geneseq/geneseq/NA1999 DAT:Z22611

seq documentation\_block:

seq\_documentation\_block;  
ID Z22611 standard; DNA; 330 BP.

ID  
XX  
AC

18-JAN-2000 (first entry)

DI 10 CAN 2000 (United Energy)  
XX Arabidopsis Na/H antiport transporter gene from EST clone T75860.  
NE

|    |  |
|----|--|
| KW | Sodium; proton; antiport; transporter; salt tolerance; salt management;    |
| KW | transgenic plant; survival; soil; farming; accumulation; irrigation;       |
| KW | crop; ss.  |
| OS |  |
| OS | <i>Arabidopsis thaliana</i> .  |
| XX |  |
| PN | WO9947679-A2.  |
| XX |  |
| PD | 23-SEP-1999.   |
| XX |  |
| PF | 18-MAR-1999; 99WO-CA00219.   |
| XX |  |
| XX | 18-MAR-1998; 98US-0078474.   |
| PR | 15-JAN-1999; 99US-0116111.   |
| PR |  |
| XX | (BLUM/) BLUMWALD E.  |
| PA | (APSE/) APSE M.  |
| PA | (SNED/) SNEDDEN W.   |
| PA | (AHAR/) AHARON G.  |
| XX |  |
| PI | Blumwald E, Apse M, Snedden W, Aharon G;                                   |
| XX |  |
| XX | WPI; 1999-571840/48.   |
| DR |  |
| PT | Nucleic acid molecules encoding sodium/proton transport polypeptides,      |
| PT | useful in genetic engineering salt tolerance in crop plants -              |
| XX |  |
| PS | Disclosure; Fig 8H; 93pp; English.   |
| XX |  |
| CC | The invention relates to an isolated nucleic acid molecule encoding        |
| CC | a plant Na/H antiport (PNHX) transporter polypeptide, or a fragment        |
| CC | and capable of increasing salt tolerance in a cell. This sequence          |
| CC | corresponds to an Na/H antiport transporter gene fragment from the         |
| CC | <i>Arabidopsis</i> expressed sequence tag (EST) library clone T75860. The  |
| CC | Na/H transporter polypeptides provide a means of intracellular salt        |
| CC | management, particularly in plants. The sequences are useful for           |
| CC | producing transgenic plants that are capable of surviving in soil          |
| CC | with high salt levels that would normally inhibit growth of the crop       |
| CC | species. This would be useful in farming land in areas that are gener-     |
| CC | ally considered unproductive through salt accumulation and poor irrigation |
| CC | e.g. in India, Australia, and prairies in USA or Canada. Commercial        |
| CC | crops, such as potatoes, tomatoes, brassica, cotton, sunflower,            |
| CC | strawberries, spinach, lettuce, rice, soybean, corn, wheat, barley,        |
| CC | amixplex, sorghum, alfalfa, salicornia and others would benefit from       |
| CC | increased salt tolerance.  |
| XX |  |
| SO | Sequence 330 BP; 92 A; 68 C; 66 G; 91 T; 13 other;                         |

```

alignment_scores:
  Quality: 415.00      Length: 85
  Ratio: 5.061        Gaps: 0
  Percent Similarity: 96.471  Percent Identity: 95.294

alignment_block:
  US-09-271-584A-2 x Z22611 ..

Align seg 1/1 to: Z22611 from: 1 to: 330

358  PheProLeuSerPheLeuSerAsnLeuAlaLysLysAsnGlnSerGluLy 374
      |||||
1    TTTCGCTTATCGTTCTTCTATCTAAGTACCAAGAGAAATCAAGCGAGAA' 50
      |||||

374  sileAsnPheAsnMetGlnValValIleTTPTrpSerGlyLeuMetArgG 391
      |||||
51  ANTCAACTTTAACATGCAGGTTGTGATTGTGGTGCTGGTCTCATGAGAG 100
      |||||

391  lyAlaValSerMetAlaLeuAlaValTrpAsnLysPheThrArgAlaGlyHis 407
      |||||
101  GTGCTGTATCTATGCTCTTGATACAAACAGTTTACAGGGCGGGCAC 150
      |||||

408  ThrAspValArgGlyAsnAlaIleMetIleThrSerThrIleThrValCy 424
      |||||

```



810 GAACCTGGAGTGTGCTCTCTTTGACTTTTCCCGTATTTTTCGTAAGTAG 859  
259 ..... 259  
860 ACAAACAACCTCCTCCTCTCTCTCGTATTTATGACAACACATTTCTTCC 909  
260 ..... 260  
910 CCCTAATGATTCTGGTATTCTCTAGATACATGCTGCAGAGGTGTT 959  
269 AspLeuSerGlyIleLeuThrValPhePheCysGlyIleVal..... 282  
|||||.....  
960 GGTCTCTCCGCATTTGTTTATCTCTTTCACAGGAATTGTAATCGCGA 1009  
282 ..... 282  
1010 GTCATTGTAGCTTTTACATCTTAGTGTGATGTTAATATCTTGGAAAGACAT 1059  
283 .....Met..SerHisTyrThrT 288  
|||.....  
1060 ATTTAGGCTGCCTAATATAGTGTCTAGTGTAGTTATGAAGCGCTACACTT 1109  
288 rpHisAsnValThrGluSerSerArgIleThrThrLysHisThrPheAla 304  
: |||.....  
1110 TCTCAAAATCTCTCAGAGCTTCACAGAGTTTCGTATCTCTTTTTCAC 1159  
305 ThrLeuSerPheLeuAlaGluThrPheIlePheLeuTyrValGlyMetAs 321  
: |||.....  
1160 TTGATATCTTCGGTAGCAGAAATCTTACCGTTTATTCATGGGATTTGA 1209  
321 p...AlaLeuAspIleAspLysTrp..... 328  
|||.....  
1210 TATTGCCATGAGCAGCATAGCTGGTCCCATGTTGGGTTTATCCTTTCT 1259  
328 ..... 328  
1260 CTATTGTATCCTCATTTACTGATGTCAGTGATTTGATGAGTGTAGTC 1309  
329 .....ArgSerValSerAs 333  
|||.....  
1310 AGTGTGTAAATCCTTGCATTTACCTTTGCTGCTGCTTTCATGACTGA 1359  
333 pThr.....ProG 336  
|||.....  
1360 CATCAGTTGTTTATTGGTGGCTAGCTGACTAAATGCTTTTATCCTG 1409  
336 lyThrSerIleAlaValSer.....SerIleLe 345  
|||.....  
1410 GCTGATCGCTTCATATCACCAGTGTTCGATTCGGATTTACCTATATG 1459  
345 uMetGlyLeuValMetValGlyArgAlaAlaPheValPheProLeuSerP 362  
: |||.....  
1460 TTCGCAATGCTTTCACCGCGGCTGTCATGTATTGGGTGTCAT 1509  
362 heLeuSerAsnLeuAlaLysLysAsnGlnSerGluLysIleAsnPheAsn 378  
: |||.....  
1510 ATTTGGTCAACCTATTAGACAG...GAGAACACAGAAATACCTATGAAG 1556  
379 MetGlnValValIleTyrTrpSerGlyLeuMetArgGlyAlaValSerMe 395  
|||.....  
1557 CACCAAAAGCCCTTTGTATAGTGGACTT...CGAGGGCAATGGCATT 1603  
395 tAlaLeuAlaTyrAsnLysPhe.....ThrArgAlaGlyHisThrAspV 410  
: |||.....  
1604 TGCATTGCATCTTCAATCATTATCATGCTACACAGAGGTCAC..... 1646  
410 alArgGlyAsnAlaIleMetIleThrSerThrIleThrValCysLeuPhe 426  
: |||.....  
1647 .....GGCCAAATCATCTTTTACTGCAACCAACATATTGTTGTTGTC 1688  
427 SerThrValValPheGlyMetLeuThrLysProLeuIleSerTyrLeu 443  
: |||.....  
1689 ACGGTTTACTAATAGGAGGTTCACAGGTAAATGTTGGAAGCTTTC.. 1736

443 uproHisGlnAsnAlaThrThrSerMetLeuSerAspAsnThrProL 460  
: |||.....GAAGTTGTAGTGACGAT..... 1754  
1737 ..... 1737  
460 ysSerIleHisIleProLeuLeuAspGlnAspSerPhe.....Ile 473  
|||||.....  
1755 .....CTTGAT...GACTCCATGTCTGAAGGCTTT 1781  
474 GluProSerGlyAsnHisAsnValProArgProAspSerIleArgGlyP 490  
|||||.....  
1782 GAAGAGAGGATCATCATGATGCTCCTCTCTTTTAGCATTTGAGCTTC 1831  
490 eleuThrArgProThr.....A 496  
: |||.....  
1832 ATCTGACGAGGATACATCATCAGGAAGCAGGTTCAAGATGAAGCTGA 1881  
496 rgThrValHisTyrTyrTrpArgGlnPhe.....AspAspSerPhe 509  
: |||.....  
1882 AGGAGTTTTCACAAAACCACTACATCATTCACGCGCTTGGACAAAACCTT 1931  
510 MetArgProValPhe 514  
: |||.....  
1932 CTGACTCCGTTCTTC 1946

seq\_name: /SIDS6/gcgdata/geneseq/geneseqn/NA1999.DAT:Z22593

seq\_documentation\_block:

ID Z22593 standard; cDNA; 714 BP.

XX

AC

XX

XX

DT

XX

DE

XX

KW

KW

KW

XX

OS

XX

PN

XX

PD

XX

XX

PF

XX

PR

XX

PR

XX

PA

XX

PA

XX

PA

XX

PI

XX

DR

XX

DR

XX

PT

XX

PS

XX

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

Nucleic acid molecules encoding sodium/proton transport polypeptides,  
useful in genetic engineering salt tolerance in crop plants -  
Claim 48; Fig 1C1; 93pp; English.

The invention relates to an isolated nucleic acid molecule encoding a  
plant Na/H antiport (PNHX) transporter polypeptide, or a fragment and  
capable of increasing salt tolerance in a cell. This sequence corresponds  
to part of the gene encoding the AtNHX3 transporter from A. thaliana.  
The Na/H transporter polypeptides provide a means of intracellular salt  
management, particularly in plants. The sequences are useful for  
producing transgenic plants that are capable of surviving in soil with

Sodium; proton; antiport; transporter; salt tolerance; salt management;  
transgenic plant; survival; soil; farming; accumulation; irrigation;  
crop; ss.

Arabidopsis thaliana.

WO9947679-A2.

23-SEP-1999.

18-MAR-1999; 99WO-CA00219.

18-MAR-1998; 98US-0078474.

15-JAN-1999; 99US-0116111.

(BLUM/) BLUMWALD E.

(APSE/) APSE M.

(SNED/) SNEDDEN W.

(AHAR/) AHARON G.

Blumwald E, Apse M, Snedden W, Aharon G;

WPI: 1999-571840/48.

P-PSDB; Y40903.





3803 TGGATTGACCAACATATTCAACTGGTTCCTACTACTGGTGTGTCTAAT 3754  
344 eLeuMetGlyLeuValMetValGlyArgAlaAlaPheValPheProLeu 361  
3753 TATPACCGCGGTGCTCTTTTTCATTCGTTCTTAGCCCATTTTCAGTGTT 3704  
361 erPheLeuSerAsnLeuAlaLysLysAsnGlnSerGluLysIleAsnPhe 377  
3703 ACACGCTTCAACATGGACTTAAGAAA.....GTTAAAAACGCAATG 3663  
378 AsnMetGlnValValIleTrpTrpSerGlyLeuMetArgGlyAlaValSe 394  
3662 AATGAATGCTGATTTAACTTTTGGCGGT...GTTAAAGGAACAGTGAG 3616  
394 rMetAlaLeuAlaTyrAsnLysPheThrArgAlaGlyHisThrAspValA 411  
3615 TTTAGCCACCATCTTTATTTTCCACACGCGATTAAACGGCGAGCATTTTA 3566  
411 rgGlyAsnAlaIleMetIleThrSerThrIleThrValCys...LeuPhe 426  
3565 AAGAAGCAGCGTTATTACTA.....TTTATTACGGCTGTGTATTATTG 3522  
427 SerThrValValPheGlyMetLeuThrLysProLeuIleSerTyrLeuLe 443  
3521 GTGACTCTGATTGGCGGATTCTAGTTTACCATTTTAA..... 3483  
443 uProHisGlnAsnAlaThrThrSerMetLeuSerAspAsnThrProL 460  
3482 .....ACTGATTCAGACGACGGAAGTGAATAAG 3452  
460 ysSerIleHisIleProLeu.....LeuAspGln 469  
3451 AGTCAGTCGAAATGATTATTATTCGTCATGTCATTGATCTTTARAACAG 3402  
470 AspSerPheIleGluProSerGlyAsnHisAsn 480  
3401 CAAATGAAGAGAAATCCTCGTGGTGAATGAAT 3369  
seq\_name: /SID56/gcgdata/geneseq/geneseq/NAL1997.DAT:V74486  
seq\_documentation\_block:  
ID V74486 standard; DNA; 8339 BP.  
XX V74486;  
XX  
XX 16-MAR-1999 (first entry)  
XX Staphylococcus aureus contig SEQ ID #175.  
XX Computer readable medium; vaccine; S aureus infection; immunodetection;  
KW cellulitis; eyelid infection; food poisoning; osteomyelitis; therapy;  
KW skin infection; surgical wound infection; scalded skin syndrome;  
XX toxic shock syndrome; ds.  
OS Staphylococcus aureus.  
XX  
XX Key Location/Qualifiers  
FH 601..660  
FT misc\_feature /\*tag= a  
FT /note= "these bases represent a line of missing text in  
FT the sequence listing in the specification. They  
FT are included to maintain the nucleotide numbering  
FT given in the specification for this DNA sequence"  
FT 2401..2460  
FT misc\_feature /\*tag= b  
FT /note= "these bases represent a line of missing text in  
FT the sequence listing in the specification. They  
FT are included to maintain the nucleotide numbering  
FT given in the specification for this DNA sequence"  
FT 4201..4260  
FT misc\_feature /\*tag= c  
FT /note= "these bases represent a line of missing text in  
FT the sequence listing in the specification. They

4504 ATATTGAGCGGTCTAATGGTTAACT..... 4575  
69 yLysSerSerHisLeuLeuValPheSerGluAspLeuPhePheIleTyrL 86  
4574 .AAATGGGCGAGGAATTAATGAGCTGAAGTGTTTTAGTAATGA 4526  
86 euLeuProIleIlePheAsnAlaGlyPheGlnValLysLysGln 102  
4525 TTATTGACCGCTTTATTATTAGAAAGGGAGGAGCTAATCAATCTCAAAGCA 4476  
103 PhePheArgAsnPheValThrIleMetLeuPheGlyAlaValGlyThrI 119  
4475 ATGATGAAAATTTTCGGGATGATCTTTCTTGGCCCTTTTGGTGCT 4426  
119 eileSerCysThrIleIleSerLeuGlyValThrGlnPhePheLysLysL 136  
4425 AATT.....ACCTTAGTGGAGTCGCTGACC.....CTGCATA 4391  
136 euAspIleGlyThrPheAspLeuGlyAspTyrLeuAlaIleGlyAlaIle 152  
4390 TGATTTTGGCTACTTTACCATAGCTAGCTGCTTTCGCTTTTGGAGCGGCT 4341  
153 PheAlaAlaThrAspSerValCysThrLeuGlnValLeuAsnGlnAspG 169  
4340 TTAGGCCAACGATGGCGGTAGGTTCGTATCAGTCAATGA 4291  
169 uThrPro...LeuLeuTyrSerLeuValPheGlyGluGlyValValAsnA 185  
4290 CATTCCTCAACGTTTCGATGAACATTTATCTGTGTGAAGGACTATTAA 4241  
185 spAlaThrSerValValPheAsnAlaIleGlnSerPheAspLeuThr 201  
4240 ATGCATCAGCGGTTACAGCCCTT.....CAATTGCAATTAGCG 4203  
202 HisLeuAsnHisGluAlaAlaPheHisLeuLeuGlyAsnPheLeuTyrL 218  
4202 GCTTTA.....TTAAGCGGCTCATTTTCGCGGT 4174  
218 u.....PheLeuLeuSerThrLeuLeuGlyAlaAlaThrG 230  
4173 GGAAGCGGCGAGGACCTTGGTTATTTCCAGCATTTGGTGGAGCGGTAGTCG 4124  
230 lyLeuIleSerAlaTyrValIleLysLysLeuTyrPheGlyArgHisSer 246  
4123 GAATCGTGTAGTCTATTAAACAGCAAAATGTT.....CATTTA 4083  
247 ThrAspArgGluValAla.....LeuMetMe 255  
4082 CTGAAACTGCAGTCGCAAAAGATCTTACAGGCTACTTGTGATTGAAT 4033  
255 tLeuMetAlaTyrLeuSerTyrMetLeuAlaGluLeuPheAspLeuSerG 272  
4032 ACTCTTGCCATTTATGCTTATTTAATTCAGAGTTGATTGGCGTTCTG 3983  
272 lyIleLeuThrValPhePheCysGlyIleValMetSer..... 284  
3982 GAATATTGCTGCTGTGATGCTGGAATATCAAGCGCTAGATTGAAA 3933  
285 .....HisTyrThrTrpHisAsnValThrGluSerSerAr 296  
3932 AAAGTTTCGCTTTTGAAGCTGAGCTGTCAAAATGTTGGCGAAAGT..... 3888  
296 gileThrThrLysHisThrPheAlaThrLeuSerPheLeuAlaGluThrP 313  
3887 .....ATTGGAGTTGATTGAGTTCACCTTTAATGCGT 3854  
313 heilePheLeuTyrValGlyMetAspAlaLeuAspIleAspLys..... 327  
3853 TAGTTTCTTATTATTAGGAATTTGAACCTTTCCTCAAGTGTTCGCGCAATT 3804  
328 TrpArgSerValSerAspThrProGlyThrSerIleAlaValSerSerI 344



```
XX Oryza sativa.
OS
XX WO9947679-A2.
PN
XX
XX 23-SEP-1999.
PD
XX
XX 18-MAR-1999; 99WO-CA00219.
PF
XX
XX 18-MAR-1998; 98US-0078474.
PR
XX 15-JAN-1999; 99US-0116111.
PR
XX
XX (BLUM/) BLUMWALD E.
PA
XX (APSE/) APSE M.
PA
XX (SNED/) SNEDDEN W.
PA
XX (AHAR/) AHARON G.
XX
XX Blumwald E, Apse M, Snedden W, Aharon G;
XX
XX WPI; 1999-571840/48.
XX
XX Nucleic acid molecules encoding sodium/proton transport polypeptides,
PT useful in genetic engineering salt tolerance in crop plants
XX
XX Disclosure; Fig 8D; 93pp; English.
XX
XX The invention relates to an isolated nucleic acid molecule encoding
CC a plant Na/H antiport (PNHX) transporter polypeptide, or a fragment
CC and capable of increasing salt tolerance in a cell. This sequence
CC corresponds to an Na/H antiport transporter gene fragment from the
CC rice expressed sequence tag (EST) library clone C91861. The Na/H
CC transporter polypeptides provide a means of intracellular salt
CC management, particularly in plants. The sequences are useful for
CC producing transgenic plants that are capable of surviving in soil
CC with high salt levels that would normally inhibit growth of the crop
CC species. This would be useful in farming land in areas that are generally
CC considered unproductive through salt accumulation and poor irrigation,
CC e.g. in India, Australia, and prairies in USA or Canada. Commercial
CC crops, such as potatoes, tomatoes, brassica, cotton, sunflower,
CC strawberries, spinach, lettuce, rice, soybean, corn, wheat, barley,
CC atriplex, sorghum, alfalfa, salicornia and others would benefit from
CC increased salt tolerance.
XX
XX Sequence 268 BP; 43 A; 55 C; 105 G; 61 T; 4 other;
SQ

alignment_scores:
Quality: 190.00 Length: 66
Ratio: 3.276 Gaps: 1
Percent Similarity: 87.879 Percent Identity: 57.576

alignment_block:
US-09-271-584A-2 x 222607 ..
Align seg 1/1 to: 222607 from: 1 to: 268

15 SerAspHisAlaSerValValAlaLeuAsnLeuPheValAlaLeuLeuCy 31
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
72 GCGGACTACGGCTCGATCGCGCGGTGGGATGTTCTGTCGCTCATCTG 121

31 sAlaCysIleValLeuGlyHisLeuLeuGluClnAsnArgTrpMetAsnG 48
|||||:::|||||:::|||||:::|||||:::|||||:::|||||
122 CGTCTGCATFCGTCGCGCACCTCTCTGAGGAGAGCCGATGGATGAACG 171

48 LuSerIleThrAlaLeuLeuIleGlyLeuGlyThrGlyValThrIleLeu 64
|||||:::|||||:::|||||:::|||||:::|||||:::|||||
172 AGTCATCATCCCGCGTATCATCTCGGTTGGTACTTGGAGGAGTGNTT 221

65 LeuIleSerLysGlyLysSerSerHisLeuLeuValPheSerGlu 79
:::|||||:::|||||:::|||||:::|||||:::|||||
222 GNATGGTGTGCGAGTGGAGACACTCGGNATACTGTTGTTGACCGAG 267

seq_name: /SID56/gcgdata/geneseq/geneseq/Nal198.DAT.V52272
```

```
seq_documentation_block:
ID_V52272 standard; DNA; 15363 BP.
XX
AC V52272;
XX
DT 23-OCT-1998 (first entry)
XX
DE Streptococcus pneumoniae genome fragment SEQ ID NO:139.
XX
XX Streptococcus pneumoniae; S. pneumoniae; genome; diagnosis; assay;
KW computer readable medium; vaccine; pharmaceutical composition; ds.
XX
XX Streptococcus pneumoniae.
OS
XX WO9818931-A2.
PN
XX 07-MAY-1998.
PD
XX
XX 30-OCT-1997; 97WO-US19588.
PF
XX
XX 31-OCT-1996; 96US-0029960.
PR
XX (HUMA-) HUMAN GENOME SCI INC.
PA
XX
XX Barash SC, Choi GH, Dillon PJ, Dougherty BA, Fannon M;
PI Kunsch CA, Rosen CA;
XX
XX WPI; 1998-272225/24.
XX
XX Computer-readable medium with recorded Streptococcus pneumoniae
PT polynucleotide sequences - useful in diagnostic kits and assays, and
PT pharmaceutical compositions and vaccines for Streptococcus
PT pneumoniae
XX
XX Claim 1; Page 931-939; 1409pp; English.
XX
XX The present invention describes a computer readable medium which has
CC the nucleotide sequences SEQ ID NO:1 to 391 (V52134 to V52524) recorded
CC on it, or a representative fragment or a sequence at least 95% identical
CC to SEQ ID NO: 1 to 391. The nucleotide sequences depicted in SEQ ID NO:1
CC to 391 (V52134 to V52524) are genomic fragments from Streptococcus
CC pneumoniae. The present invention also describes an isolated nucleic acid
CC molecule encoding a homologue of any of the fragments of the S.pneumoniae
CC genome (SEQ ID NO:1 to 391) where the nucleic acid molecule is produced
CC by a process comprising: (a) screening a genomic DNA library using as a
CC probe a target sequence defined by any of the sequences in SEQ ID NO:1
CC to 391, identifying members of the library which contain sequences
CC that hybridise to the target sequence and isolating the nucleic acid
CC molecules from the members; or (b) isolating mRNA, DNA or cDNA produced
CC from an organism, amplifying nucleic acid molecules whose nucleotide
CC sequence is homologous to amplification primers derived from the
CC fragment of the S. pneumoniae genome to prime the amplification and
CC isolating the amplified sequences. The computer readable medium can be
CC used in a computer-based system for identifying fragments of the
CC S. pneumoniae genome of commercial importance, or expression modulating
CC fragments of the S. pneumoniae genome. Products from the present
CC invention can be used in diagnosis kits and assays, and pharmaceutical
CC compositions and vaccines for S. pneumoniae.
XX
XX Sequence 15363 BP; 4357 A; 3360 C; 2915 G; 4731 T; 0 other;
SQ

alignment_scores:
Quality: 185.50 Length: 486
Ratio: 0.733 Gaps: 21
Percent Similarity: 52.058 Percent Identity: 20.576

alignment_block:
US-09-271-584A-2 x V52272/rev ..
Align seg 1/1 to reverse of: V52272 from: 1 to: 15363
```

20 ValValAlaLeuAsnLeuPheValAlaLeuLeuCysAlaCysIleValLe 36  
2383 CTGATTTACCTCATCTATTTTACTGGTTTGTGATGCTCAAGTACAC 2334  
36 uGlyHisLeuLeuGluGluAsnArgTrpMetAsnGluSerIleThrAlaL 53  
2333 CAATAAGCTCCTG.....CCCTTTTGGCCTCTCC 2305  
53 euleuileGlyLeuGlyThrGlyValThrIleLeuLeuIleSerLysGly 69  
2304 CTTTGGTGAATTCCTTTGGGAATGTGATGGCTCTCTTTTACCACAT 2255  
70 LysSerSerHisLeuLeuValPheSerGluAspLeuPhePheIleTyrIle 86  
2254 ACTGACTTTTCACCTT.....AATACGGAGTTGTTTTCGCACCTGGT 2214  
86 uleuProPheIlePheAsnAlaGlyPheGlnValLysLysGlnP 103  
2213 TATCGGACCTCTGCTTTCCGAGAGCTGAAGAAGCAGATGTACGGCTA 2164  
103 hepheArgAsnPhe.....ValThrIleMetLeuPheGlyAlaValGly 117  
2163 TTTTAAACACTGGGAATCATGTTATCTCATATTTCCAGTGATT... 2117  
118 ThrIleIleSerCysThrIleIleSerLeuGlyValThrGlnPhePhe 134  
2116 .....TTTATCTC 2109  
134 stysLeuAspIleGly.....ThrPheAspL 143  
2108 GACCCCTGAGTTGGTGGCTTGGCCCATCTCTTTGGTTTCAGCCTTCCT 2059  
143 euGlyAspTyrLeuAlaIleGlyAlaIlePheAlaAlaThrAspSerVal 159  
2058 TGGCAGCTTGCCTGGCTGTGGGGAGCCCTTGGTCTACGGAGCTGGTG 2009  
160 CysThrLeuGlnValLeuAsnGlnAspGluThrProLeuLeuTyrSer... 175  
2008 GCCTTTCCTCTCTCTTCGGAGCGTTTAGCTTTCTAGCGCGTGTCCA 1959  
176 LeuValPheGlyGluGlyValValAsnAspAlaThrSerValValP 192  
1958 TATTCTTAAGGGCGAAGACTCTTGAATGATGCTTCTGGTTGGTGCTT 1909  
192 heAsn.....AlaIleGlnSerPheAspLeuThrHis 202  
1908 TTCAGGTAGCTTTGACAGCTTGGCAACTGGAGCTTTTCTCTGGGGCAA 1859  
203 LeuAsnHisGluAlaAlaPheHisLeuLeuGlyAsnPheLeuTyrLeuPh 219  
1858 GCTAGCAGTTCCGCTCATCTTTTCATCTCCTAGCGGGTTTT... 1820  
219 eLeuLeuSerThrLeuLeuGlyAlaAlaThrGlyLeuIleSerAlaTyrV 236  
1819 .....TTAATTGGATTTTAAACAGCCATGACCAACCGCTTCC 1783  
236 alIleLysLysLeuTyrPheGlyArgHisSerThrAsp.....ArgGlu 250  
1782 TCCATACCTTCTGTGTAAGTGGCCG...GCAACGGATATTGCCAGTGA 1736  
251 ValAlaLeuMetMetLeuMetAlaTyrLeuSerTyrMetLeuAlaGluLe 267  
1735 CTTTATTAGATTGAGTTGGCTCTAGTGACCTTCTTCTGGCAGAAGA 1686  
267 uPheAspLeuSerGlyIleLeuValPhePheCysGlyIleVal...M 283  
1685 AGTTTCATGTTTCAGGTATTATTGCGCTCGTAGTGTGGAATTTTAAAG 1636  
283 etSerHisTyrThrTrpHisAsnValThrGluSerSerArgIleThr... 298  
1635 CAAGTCGCTTCAAGAAATTCACGCTCTCCGAAGCCCAAGTGGATACGGTG 1586  
299 ThrLysHisThrPheAlaThrLeuSerPheLeuAlaGluThrPheIlePh 315

11585 ACCGAGCGGTCTGGCATACAGTACCTTTATGTCTCAACGGTTCTGTCTT 1536  
315 eLeuTyrValGlyMetAspAlaLeuAspIleAspLysTrpArgSerVals 332  
1535 TGTGATTTTAGGATGGATTG.....GAAATGATAG 1504  
332 erAspThrProGlyThrSerIleAlaValSerSerIleLeuMetGlyLeu 348  
1503 CAGAACCTTATCTTGACCAATCAATCATATAATCTCTACTTTTATTGCTA 1454  
349 ValMetValGlyArgAlaAlaPheValPheProLeuSerPheLeuSer... 364  
1453 TCTCTATGCGCCCTACCTTTTCTCTTTTGTCTTTCATTCGTTTATTATGAT 1404  
365 .....AsnLeuAlaLysLysAsnGlnSerGluLysIleAsnP 377  
1403 CTATGGCTATTATGCTATAGAACCGCGCTCAAGAAAAAGCTAAATA 1354  
377 heAsnMetGln.....ValValIleTrpTrpSerGlyLeuMetArgGly 391  
1353 AGTATATGAAGGACATGTTCTCTTGACCTTTTCAGGT...GTTAAGGGA 1307  
392 AlaValSerMetAlaLeuAlaTyrAsnLysPheThrArgAlaGlyHisTh 408  
1306 ACGGTGTCGATTGCT..... 1292  
408 rAspValArgGlyAsnAlaIleMetIleThrSerThrIleThrValCysL 425  
1291 .....ACGATTCCTTGTATACCAAGTAATAGAACAG...G 1258  
425 euPheSerThrValPheGlyMetLeuThrLysProLeuIleSerTyr 441  
1257 AGTATCTCTCTTTTGTCTTCTGTCAGGTGTGAGCTTGTAGCTTT 1208  
442 Leu.....LeuProHisGlnAsnAlaThrThrSerMetLe 453  
1207 TTAACAGTCTCTTGGTCTTGCCTCAT.....CT 1179  
453 uSerAspAspAsnThrProLysSer.....IleHisIleProLeuL 467  
1178 TTCTGTAGCAGAGGAGGAAGCAAGGATTATCTTATGCATATCCCATTT 1129  
467 euAspGln 469  
1128 TGAATGAA 1121  
seq\_name: /SIDS6/gcgdata/geneseq/geneseq/NA1997.DAT.V74657  
seq\_documentation\_block:  
ID V74657 standard; DNA; 1294 BP.  
XX V74657;  
XX AC  
XX DT  
XX 16-MAR-1999 (first entry)  
XX Staphylococcus aureus contig SEQ ID #346.  
DE Computer readable medium; vaccine; S.aureus infection; immunodetection;  
XX cellulitis; eyelid infection; food poisoning; osteomyelitis; therapy;  
KW skin infection; surgical wound infection; scalded skin syndrome;  
XX toxic shock syndrome; ds.  
XX Staphylococcus aureus.  
OS  
XX EP786519-A2.  
XX  
XX 30-JUL-1997.  
XX  
XX 07-JAN-1997; 97EP-0100117.  
XX  
XX 05-JAN-1996; 96US-0009861.  
PR

139 ythrPheaspLeuGlyAspTyrLeuAlaIleGlyAlaIlePheAlaAat 150  
568 TGGCTTTACCAATAGCAGCGCTTTTGCAATAGCAGCAATTTATGTGCCTTA 519  
156 hrAspSerValCysThrLeuGlnValLeuAsnGlnAspGluThrProLeu 172  
|||||:|||||: : : : : ||||  
518 CTGTCGACGTAGCAGTACTGCTATTACACGCGAAAAATTATTACCTAAA 469  
173 ...LeuTyrSerLeuValPheGlytGlUGLyValValAsnAspAlaThrSe 188  
: : : : : |||||: : : : : |||||: : : : :  
468 GGTTCCTCATCACCATTTTAGAAGGTGAATCTTTACTCAATCATGCAGCAGG 419  
188 rValValValPheAsnAlaIleGlnSerPheAspLeuHisLeuAsnH 205  
: : : : : |||||: : : : : |||||:  
418 TATCATTTCAITTAATAATTGCTGTTACTGCTAGTAGTACT... 379  
205 iSGLUALAAlaPheHisLeuLeuGlyAsnPheLeuTyrLeuPheLeuLeu 221  
: : : : : |||||: : : : : |||||: : : : :  
378 ....GGTACCCTTTTCATATTCT...CAAGCTGTGAACAATTTATTCATT 337  
222 SerThrLeuLeuGlyAlaAlaThrGlyLeuLeuSerAlaTyrValIleY 238  
|||||:|||||: : : : : |||||: : : : : |||||: : : : :  
336 TCACAACATACTAGGGTGACTTATTGGTGCAATTATGGATTCGTTGCTGT 287  
238 sLysLeuTyrPhe.....GlyArgHisSerThrAspA 249  
: : : : : |||||: : : : :  
286 TAGAATACGTATAGACTTAACCGCTAATAAAGGTTTTAAAGAATAACAATA 237  
249 rgGluValAlaLeuMetMetLeuMetAlaTyrLeuSerTyrMetLeuAla 265  
236 CCTAACITTTTATTCAAATTAATAACTCCTTTTTCGTGTACTTTTATGCA 187  
266 GluLeuPheAspLeuSerGlyLeuLeuThrValPhePheCysGlyIleVa 282  
||| |||||:|||||: : : : : |||||: : : : :  
186 GAAGAAGTTCATGCTTCAGGTATCATTCGACTCTTAATCCAGTCTAAT 137  
282 lMetSerHisTyrThrTrpHisAsnValThrGluSerSerArgIleThrT 299  
: : : : : |||||: : : : : |||||: : : : :  
136 C.....CATGGTTTA...GAACGCGACCGTTTAATTA 108  
299 hrlYshisThr.....PheAlaThrLeu 306  
: : : : : |||||: : : : :  
107 GAGCAAACGGAACATAAATGAATTACCATCAATTTTGGACACATTT 58  
307 SerPheLeuAlaGluThrPheIlePheLeuTyrValGly 319  
57 AGTTATGCCCTTAATAGTTTGTGTTATTTGTTGTTATTTAGGT 19

282 lMetSerHisTyrThrTrpHisAsnValThrGluSerSerArgIleThrTT 299  
:  
: |||::|||:: ||| :|||::  
136 C.....CATGTTTA...GAACGGACCGCTTAAATTA 108  
  
299 hrlYshisThr.....PheAlaIthrLeu 306  
:||||  
107 GAGCAACAACGGAACTCAAATGAATTACCATCAATTTGGACACACTTT 58  
  
307 SerPheLeuAlaGluThrPheIlePheLeuTyrValGly 319  
|||::: |||::|||::: |||::  
57 AGTATGCCCTTAAATGGTTTGTGATTGTTGTTGATTAGT 19







```
Database: Issued_Patents_NA:
Database sequences: 280836
Database length: 80580151
Search time (sec): 75.940000
```

[illegible]

53 euLeuIlegIyLeugIvThrgIvvaIThrIeLeuIeIeSerIvsGlyv 69





```
386 rGlyLeuMetArgGlyAlaValSerMet.....AlaLeuAlaTyrAsnL 401
|||||: : : : : : : : : : : : : : : : : : : : : : : :
7392 AGGTTGTCATCGCGAGTATATTGACTTTTACTTATGACACTTATCATGATAA 7441
|||||: : : : : : : : : : : : : : : : : : : : : : : :
401 ysPheThrArgAlaGlyHisThrAspValArgGly..... 412
||| ||| |||: : : : : : : : : : : : : : : : : : : :
7442 AAGAAACATCTTGGGAAATTAACAATATAGAAAAATTTAAACGTAACAA 7491
413 .....AsnAlaIleMetIleThrSerTh 420
|||||: : : : : : : : : : : : : : : : : : : : : : : :
7492 ATACATGAACCATGGCTATTAGTTTACCGCTGTGATTTTAAATGTTACT 7541
420 rIleThrValCysLeuPheSerThrValValPheGly 432
||| ||| |||: : : : : : : : : : : : : : : : : : : :
7542 CATTCAGTTATCTCTTCTTCTTCAACGTTTTTGGC 7578

seq_name: /cqn2.6/ptodata/1/ina/5A_COMB.seq:us-08-920-827-6

seq_documentation_block:
; Sequence 6, Application US/08920827
; Patent No. 5770375
; GENERAL INFORMATION:
; APPLICANT: Ohno, Tsuneya
; APPLICANT: Matsuhisa, Akio
; APPLICANT: Uehara, Hirotugu
; APPLICANT: Eda, Soji
; TITLE OF INVENTION: Probe for Diagnosing Infectious Disease
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESS: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/920,827
; FILING DATE: 29-AUG-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/362,577
; FILING DATE: 27-MAR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Rin-Laures, Li-Hsien
; REGISTRATION NUMBER: 33,547
; REFERENCE/DOCKET NUMBER: 19036/32420
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312/474-6300
; TELEFAX: 312/474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8654 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: Genomic DNA
; ORIGINAL SOURCE:
; ORGANISM: Staphylococcus epidermidis
; STRAIN: Clinical isolate SE-22
; US-08-920-827-6
```

```
alignment_scores:
  Quality: 132.00      Length: 546
  Ratio: 0.557        Caps: 25
  Percent Similarity: 43.407  Percent Identity: 20.879
```

```
alignment_block:
US-09-271-584A-2 x US-08-920-827-6
Align seg 1/1 to: US-08-920-827-6 from: 1 to: 8654

1 MetLeuAspSerLeuValSerLysLeuProSerLeuSerThrSerAspHi 17
|||||: : : : : : : : : : : : : : : : : : : : : : : :
6107 ATGGCGACTAATTTATTAGTCATGCTCATAGTTTATTTCACCTCTGAGTCA 6156

17 sLaSerValValAlaLeuAsnLeuPheValAlaLeuLeuCysAlaCysI 34
|||||: : : : : : : : : : : : : : : : : : : : : : : :
6157 TCGTCAACATAAGAAAGGTTCGGGCTATGTTGCATTAATAGTCTCCTATTG 6206

34 leValLeuGlyHisLeuLeuGlu.....AsnArg 44
|||||: : : : : : : : : : : : : : : : : : : : : : : :
6207 TGACATCTACATATTTTATTATGAAAAATACCAGATGTCATTCGAAATAAG 6256

45 .....TrpMet.....AsnGluSerIleTh 51
|||||: : : : : : : : : : : : : : : : : : : : : : : :
6257 TTTATTGCTGTTTCGATTACCATGGATGCTTCAATTGATATTAATTAGA 6306

51 rAlaLeuLeuIleGlyLeuGlyThr.....GlyValThrIleLeuLeuI 66
|||||: : : : : : : : : : : : : : : : : : : : : : : :
6307 TTTAAGATTAGATGTTTAAAGTTTAAATGTTTCGGCTTAATTTATTCGCTAA 6356

66 leSerLysGly.....LysSerSerHisLeuLeuValPheSerGlu 79
|||||: : : : : : : : : : : : : : : : : : : : : : : :
6357 TAGGTTGGGTGTTATTTTATGCTACGCAATATTATCCACAGTACG 6406

80 Asp.....LeuPheIleTyrLeuLeuProIleIleIlePheAs 93
|||||: : : : : : : : : : : : : : : : : : : : : : : :
6407 GACAATCTCTCTAGATTTTTCATCTATTACTA..... 6439

93 nAlaGlyPheGlnValLysLysGlnPhePheArgAsnPheValThrI 110
6440 .....T 6440

110 leMetLeuPheGlyAlaValGlyThrIleIleSerCysThrIleIleSer 126
|||||: : : : : : : : : : : : : : : : : : : : : : : :
6441 TATTATGTTTCAGTATGATTCGATTGTAATAGCTAATAATACCCTCTA 6490

127 LeuGlyValThrGlnPhePheLysLysLeuAspIleGlyThrPheAspLe 143
|||||: : : : : : : : : : : : : : : : : : : : : : : :
6491 ATGTATGTA.....TTTGGGAACACTCACAAGTATTTCCTCATTCTGCT 6534

143 uGlyAspTyr..... 146
|||||: : : : : : : : : : : : : : : : : : : : : : : :
6535 TATATCTTATGTTGTAACAATAATGGTGAAGTCAATTAGCGCCATTCAAT 6584

147 .....LeuAlaIleGlyAlaIlePheAlaAlaThrAspSerVal 159
|||||: : : : : : : : : : : : : : : : : : : : : : : :
6585 CTTTCATGATTACAGTGTTCGTTGGCTAGCGTTATTAAACAGGATTTATC 6634

160 CysThrLeuGlnValLeuAsnGlnAspGluThrProLeuLeuTyrSerLe 176
|||||: : : : : : : : : : : : : : : : : : : : : : : :
6635 .....ATTTTATAT...AT 6645

176 uValPheGlyGluGlyValValAsnAspAlaThrSerValValPheA 193
|||||: : : : : : : : : : : : : : : : : : : : : : : :
6646 CATTACAGGAACAACAACAAATTACTGAT.....ATCTTTAATCA 6683

193 snAlaIleGlnSerPheAspLeuThrHisLeu..... 203
|||||: : : : : : : : : : : : : : : : : : : : : : : :
6684 AGCGCAATGCAATTTCCAGCACATCCTTTATTATACCAATGATTTTGTATGC 6733

204 AsnHisGluAlaAlaPheHisLeuLeuGlyAsnPheLeuTyrLeuPhe... 219
|||||: : : : : : : : : : : : : : : : : : : : : : : :
6734 TATTATTAGTGTCTTTTACCAAACTCTGCACAAATTTCCGTTTCATATTGG 6783

220 Leu.....LeuSerThrLeuLeuGlyAl 227
|||||: : : : : : : : : : : : : : : : : : : : : : : :
6784 TTACCAAGGCCATGGCAGCACCTACACCAGTAAGTCTTATCTTCATTC 6833
```

```
227 aalaThr..... 229
:|||||
6834 GGCACAAATGCTAAAGCTGAATCTTTTACATATTAGATTACACCTT 6883
230 .....GlyLeuIleSerAlaThrValIleLysLysLeuThrPheGlyArg 244
||||| :|||:|
6884 TATTGGGACTAGTAATGTTTATATTATACAGTGCACATTTGTTGTT 6930
245 HisSerThrAspArgGluValAlaLeuMetMetLeuMetAlaThrLeuSe 261
:||||| :|||:|
6931 .....CTAATAACTATGTTATTGGATC 6953
261 rTyrMetLeuAlaGluLeuPheAspSerGlyIleLeu.....T 275
:||||| :|||:|
6954 TTTAACTGCTTAGCACAATACACACTAAAGGTATACCTCGTATTCTTA 7003
275 hrValPhePheCysGlyIleValMetSer.....His 285
:||||| :|||:|
7004 CAATAAGTCAATTAGGTATGATGATTATGACATGCTAGTCTAGTGGCGT 7053
286 TyrThrTrpHisAsnValThrGluSerSerArgIleThrThrLysHisTh 302
:||||| :|||:|
7054 TAGCTCAGCACACATCAGATGATGCTCTAAGTTTATATTATTAGTTT 7103
302 rPheAlaThrLeuSerPheLeuAlaGluThrPheIlePhe.....L 316
:||||| :|||:|
7104 ATTTGCTGGCTTATTCATTAATGATCATGCGGTTTAAATGTGCAT 7153
316 euTyrValGlyMetAspAlaLeuAsp..... 324
:||||| :|||:|
7154 TATTATGGCGTGTGATGATGATGATGATGATGATGATGATGATGATGAT 7203
325 .....IleAspLysTrpArgSerValSerAspThrProGlyThrSeril 339
:||||| :|||:|
7204 CGTTTCTAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 7241
339 eAlaValSerSerIleLeuMetGlyLeuValMetValGly.....A 353
:||||| :|||:|
7242 GCATATTGTCATGCTGCTGCATATCTATGCGAGGTGTCCTTTT 7291
353 rGlaAlaPheValPheProLeuSerPheLeuSerAsnLeuAlaLysLys 369
:||||| :|||:|
7292 TAAATGGCTTTTAAAGTAAGGAATGTTTATGATTCGTTAACTAAAGCA 7341
370 AsnGlnSerGluLysIleAsnPheAsnMetGlnValValIleThrTrpSe 386
:||||| :|||:|
7342 AACGAACCTGATCAATATGCTGCTGATTAACGTTTGTGATTATTCAAT 7391
386 rGlyLeuMetArgGlyAlaValSerMet.....AlaLeuAlaTyrAsnL 401
:||||| :|||:|
7392 AGGTGTCATCGCGAGTATATTGACTTTTACTTATGACATTTACATGATA 7441
401 yPheThrArgAlaGlyHisThrAspValArgGly..... 412
:||||| :|||:|
7442 AGAAACATCTCTGGGAAATACAAATATAGAAATTTAAACGTAACAA 7491
413 .....AsnAlaIleMetIleThrSerTh 7420
7492 ATACATGAACCTGCTATTATTAGTTTACCAGCTGTGATTTTATGTTACT 7541
420 rIleThrValCysLeuPheSerThrValValPheGly 432
:||||| :|||:|
7542 CATTCAGTATCTCTTGTGTTCCAAACGTTTGTGTC 7578
```

seq\_name: /cgn2\_6/ptodata/1/ina/5A\_COMB.seq:us-08-921-177-6

seq\_documentation\_block:

; Sequence 6, Application US/08921177

; Patent No. 5798211

; GENERAL INFORMATION:

; APPLICANT: Ohno, Tsuneya

; APPLICANT: Matsuhisa, Akio

```
; APPLICANT: Uehara, Hirotsugu
; APPLICANT: Eda, Soji
; TITLE OF INVENTION: Probe for Diagnosing Infectious Disease
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/921,177
; FILING DATE: 29-AUG-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/362,577
; FILING DATE: 27-MAR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Rin-Laures, Li-Hsien
; REGISTRATION NUMBER: 33,547
; REFERENCE/DOCKET NUMBER: 19036/32420
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312/474-6300
; TELEFAX: 312/474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8654 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: Genomic DNA
; ORIGINAL SOURCE:
; ORGANISM: Staphylococcus epidermidis
; STRAIN: Clinical isolate SE-22
; US-08-921-177-6
```

alignment\_scores:

Quality: 132.00 Length: 546

Ratio: 0.557 Gaps: 25

Percent Similarity: 43.407 Percent Identity: 20.879

alignment\_block:

US-09-271-584A-2 x US-08-921-177-6 ..

Align seg 1/1 to: US-08-921-177-6 from: 1 to: 8654

```
1 MetLeuAspSerLeuValSerLysLeuProSerLeuSerThrSerAspHi 17
||| :|||:|
6107 ATGGCGACTAATTTAGTACTGCTATAGTCTTTTATTCACCTCGAGTCA 6156
17 sAlaSerValAlaLeuAsnLeuPheValAlaLeuLeuCysAlaCysI 34
| :|||:|
6157 TCGTCAACTAAGAAAGGTTGCGGCTATGTTGCATTAATAGCTCCTATTG 6206
34 leValLeuGlyHisLeuLeuGlu.....AsnArg 44
:||||| :|||:|
6207 TGACATCTACATATTTTATTATGAAATACACAGATGATTCGAAATAAG 6256
45 .....TTPMet.....AsnGluSerIleTh 51
||||| :|||:|
6257 TTTATTGCTGTCGATTACCATGGATGCTTCATTTGATTAATTTAGA 6306
51 rAlaLeuLeuIleGlyLeuGlyThr.....GlyValThrIleLeuLeuI 66
||| :|||:|
6307 TTTAAGATTAGATGTTTAACTTTAATGTTTCGGCTTAATTTATTCGCTAA 6356
```

```
66 leSerLysGly.....LysSerSerHisLeuLeuValPheSerGlu 79
||||| ||| :||| :||| :|||
6357 TAGGTGGGTGATATTTTATGCTACGCAATATTTATCCACAGTACG 6406
||||| :||| :||| :||| :|||
80 Asp.....LeuPhePheIleTyrLeuLeuProIleIlePheAs 93
||||| ||| :||| :||| :|||
6407 GACAATCTTCAGATTTTCATCTATTTACTA..... 6439
||||| :||| :||| :||| :|||
93 nAlaGlyPheGlnValLysLysLysGlnPhePheArgAsnPheValThrI 110
||||| :||| :||| :||| :|||
6440 .....T 6440
110 leMetLeuPheGlyAlaValGlyThrIleIleSerCysThrIleIleSer 126
||||| :||| :||| :||| :|||
6441 TATTTATGTTGAGTATGATGCTGATGCTGATGCTGATGCTGATGCTGAT 6490
||||| :||| :||| :||| :|||
127 LeuGlyValThrGlnPhePheLysLysLeuAspIleGlyThrPheAspLe 143
||||| :||| :||| :||| :|||
6491 ATGATATGA.....TTTGGGAAGTCAAGATATTTCCATCTTGTCT 6534
||||| :||| :||| :||| :|||
143 uGlyAspTyr..... 146
||||| :||| :||| :||| :|||
6535 TATATCTTATTTGGTACAATAATGTTGAAGTCAATAGGCCCAATCAAT 6584
||||| :||| :||| :||| :|||
147 .....LeuAlaIleGlyAlaIlePheAlaIleThrAspSerVal 159
||||| :||| :||| :||| :|||
6585 CTTTCATGATTTACAGTGTGGTGGGCTGAGCTGATTAACAGGATTTATC 6634
||||| :||| :||| :||| :|||
160 CysThrLeuGlnValLeuAsnGlnAspGluThrProLeuLeuTyrSerLe 176
||||| :||| :||| :||| :|||
6635 .....ATTATATAT...AT 6645
||||| :||| :||| :||| :|||
176 uValPheGlyGluGlyValValAsnAspAlaThrSerValValValPheA 193
||||| :||| :||| :||| :|||
6646 CATTACAGAAACAACAATTAAGTAT.....ATCTTAATCA 6683
||||| :||| :||| :||| :|||
193 snAlaIleGlnSerPheAspLeuThrHisLeu..... 203
||||| :||| :||| :||| :|||
6684 AGCAATGCAATTTCCAGACATCTTTATTTATACCAATGATTTTGATGC 6733
||||| :||| :||| :||| :|||
204 AsnHisGluAlaAlaPheHisLeuLeuGlyAsnPheLeuTyrLeuPhe.. 219
||||| :||| :||| :||| :|||
6734 TATTATTAGTGTCTTTTACCAAAATCTGCACAAATTTCCGTTTCATATTTGG 6783
||||| :||| :||| :||| :|||
220 Leu.....LeuSerThrLeuLeuGlyAl 227
||||| :||| :||| :||| :|||
6784 TTACCAAAAGGCATGGCAGCACCTACACAGTAAGTCTTATCTTCATTC 6833
||||| :||| :||| :||| :|||
227 aAlaThr.....CTAATAACTATGTTATTTGGATC 6953
||||| :||| :||| :||| :|||
6834 GCACAATGTTAAAGCGTGGATCTTTTACTATTATAGATTACACCTT 6883
||||| :||| :||| :||| :|||
230 .....GlyLeuIleSerAlaTyrValIleLysLysLeuTyrPheGlyArg 244
||||| :||| :||| :||| :|||
6884 TATTGGGACTTAGTAATGTTTATATTTATACAGTGACATTTTGTGTT... 6930
||||| :||| :||| :||| :|||
245 HisSerThrAspArgGluValAlaLeuMetMetLeuMetAlaTyrLeuSe 261
||||| :||| :||| :||| :|||
6931 .....CTAATAACTATGTTATTTGGATC 6953
||||| :||| :||| :||| :|||
261 rTyrMetLeuAlaGluLeuPheAspLeuSerGlyIleLeu.....T 275
||||| :||| :||| :||| :|||
6954 TTTAACTGTTTCCAGCAATAGAGCTTAAAGGTATACCTCGCTTATCTA 7003
||||| :||| :||| :||| :|||
275 hrValPhePheCysGlyIleValMetSer.....His 285
||||| :||| :||| :||| :|||
7004 CAATAAGTCAATGATGATGATGATGATGATGATGATGATGATGATGAT 7053
||||| :||| :||| :||| :|||
286 TyrThrTriPheHisAsnValThrGluSerArgIleThrThrLysHisTh 302
||||| :||| :||| :||| :|||
7054 TAGCTCAGCACACATCAGATGAATGCTGAAGTTTATATTTTACTTTT 7103
```

```
302 rPheAlaThrLeuSerPheLeuAlaGluThrPheIlePhe.....L 316
||||| ||| :||| :||| :|||
7104 ATTGCTGGCTTATTCATTATGAATCATCGGTTTTTAATATGTCAT 7153
||||| :||| :||| :||| :|||
316 euTyrValGlyMetAspAlaLeuAsp..... 324
||||| :||| :||| :||| :|||
7154 TATTTATGGCGTGTGATCATGATCAGGAGTCCGGAACAGCTGATATT 7203
||||| :||| :||| :||| :|||
325 .....IleAspLysTrpArgSerValSerAspThrProGlyThrSerI 339
||||| :||| :||| :||| :|||
7204 CGTTTGCTAAATGGTATGCTGAAGTCTCC.....CCTAAAT 7241
||||| :||| :||| :||| :|||
339 eaLavalSerSerIleLeuMetGlyLeuValMetValGly.....A 353
||||| :||| :||| :||| :|||
7242 GCATATGTCATGCTGCTCGTCATTTATCTATGCGAGCTGTCCTTTT 7291
||||| :||| :||| :||| :|||
353 rgAlaAlaPheValPheProLeuSerPheLeuSerAsnLeuAlaLysLys 369
||||| :||| :||| :||| :|||
7292 TAAATGGCTTTTAAAGTAAGGAAATGTTTTAGATTCTGTTAAAGCA 7341
||||| :||| :||| :||| :|||
370 AsnGlnSerGluLysIleAsnPheAsnMetGlnValValIleTrpTrpSe 386
||||| :||| :||| :||| :|||
7342 AAGCACTTGAACAATATGCTTCGTTTAAACGTTTGTGATTTTCAAT 7391
||||| :||| :||| :||| :|||
386 rGlyLeuMetArgGlyAlaValSerMet.....AlaLeuAlaTyrAsnL 401
||||| :||| :||| :||| :|||
7392 AGGTGTCATCGCGAGTATATGACTTTTACTTTATGCACTTTACATGATA 7441
||||| :||| :||| :||| :|||
401 ysPheThrArgAlaGlyHisThrAspValArgGly..... 412
||||| :||| :||| :||| :|||
7442 AAGAAACATTTCTGGGAAATTTACAATATAGAAAAATTTAAACGTAAACAA 7491
||||| :||| :||| :||| :|||
413 .....AsnAlaIleMetIleThrSerTh 420
||||| :||| :||| :||| :|||
7492 ATACATGAACCATGGCTATTATTACGAGCTGTCATTTTAAATGTTACT 7541
||||| :||| :||| :||| :|||
420 rIleThrValCysLeuPheSerThrValValPheGly 432
||||| :||| :||| :||| :|||
7542 CATTCAGATTCTCTCTTTGTTCCAAACGTTTTTGGC 7578
||||| :||| :||| :||| :|||
seq_name: /cgn2_6/ptodata/1/ina/5A_COMB.seq:US-08-362-577C-6
seq_documentation_block:
; Sequence 6, Application US/08362577C
; Patent No. 5807673
; GENERAL INFORMATION:
; APPLICANT: Ohno, Tsuneya
; APPLICANT: Matsuhisa, Akio
; APPLICANT: Uehara, Hirotsugu
; APPLICANT: Eda, Soji
; TITLE OF INVENTION: Probe for Diagnosing Infectious Disease
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/362,577C
; FILING DATE: 27-MAR-1995
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Rin-Laures, Li-Hsien
; REGISTRATION NUMBER: 33,547
; REFERENCE/DOCKET NUMBER: 19036/32420
; TELECOMMUNICATION INFORMATION:
```



```

401 yspPheThrArgAlaGlyHisThrAspValArgGly..... 412
|| ||| |||::: :::::
7442 AAGAAACATTTCTGGCGAAATTACAATAATAGAAAATTTAAACGTAAACAA 7491
413 ..... AsnAlaIleMetIleIleThrSerTh 420
||||:::::
7492 ATACATGAACCATGCTATTTAGTTTACCAGCTGTGATTTTAATGTTACT 7541
420 rIleThrValCysLeuPheSerThrValValPheGly 432
||| ||| ::::| |
7542 CATTCAGTGATCTCTTTGTTGCCAACGTTTTTGGC 7578

seq_name: /cgn2_6/ptodata/1/ina/5B_COMB.seq:US-08-920-828-6

seq_documentation_block:
; Sequence 6, Application US/08920828
; Patent No. 5853998
; GENERAL INFORMATION:
; APPLICANT: Ohno, Tsuneya
; APPLICANT: Matsuhisa, Akio
; APPLICANT: Uehara, Hirotsugu
; APPLICANT: Eda, Soji
; TITLE OF INVENTION: Probe for Diagnosing Infectious Diseases
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/920,828
; FILING DATE: 29-AUG-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/362,577
; FILING DATE: 27-MAR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Rin-Laures, Li-Hsien
; REGISTRATION NUMBER: 33,547
; REFERENCE/DOCKET NUMBER: 19036/32420
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312/474-6300
; TELEFAX: 312/474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8654 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: Genomic DNA
; ORIGINAL SOURCE:
; ORGANISM: Staphylococcus epidermidis
; STRAIN: Clinical Isolate SF-22
; US-08-920-828-6

alignment_scores:
Quality: 132.00 Length: 546
Ratio: 0.957 Gaps: 25
Percent Similarity: 43.407 Percent Identity: 20.879

alignment_block:
US-09-271-584A-2 x US-08-920-828-6 ..
Align seg 1/1 to: US-08-920-828-6 from: 1 to: 8654
```



```
230 .....GlyLeuIleSerAlaTyrValIleLysLysLeuTyrPheGlyArg 244
||||| : : : : : : : : : : : : : : : : : : : : : : : : : : :
6884 TATTGGGACTAGTAATGTTTATATATACAGCAGACATTTGTGGT... 6930
: : : : : : : : : : : : : : : : : : : : : : : : : : :
245 HisSerThrAspArgGluValAlaLeuMetMetLeuMetAlaTyrLeuSe 261
: : : : : : : : : : : : : : : : : : : : : : : : : : :
6931 .....CTAATAACTATGTTATTGGATC 6953
: : : : : : : : : : : : : : : : : : : : : : : : : : :
261 rTyrMetLeuAlaGluLeuPheAspLeuSerGlyIleLeu.....T 275
: : : : : : : : : : : : : : : : : : : : : : : : : : :
6954 TTTAACTGCTTTACGACAATACGAGTAAAAGGTATACGCTTATCTA 7003
: : : : : : : : : : : : : : : : : : : : : : : : : : :
275 hrValPhePheCysGlyIleValMetSer.....His 285
: : : : : : : : : : : : : : : : : : : : : : : : : : :
7004 CAATAGTCAATTAGTATGATTATGACAATGGTAGGTCTAGGTGGCGGT 7053
: : : : : : : : : : : : : : : : : : : : : : : : : : :
286 TyrThrTrpHisAsnValThrGluSerSerArgIleThrThrLysHisTh 302
: : : : : : : : : : : : : : : : : : : : : : : : : : :
7054 TATGCTCAGCACACATCAGATGAATGTCTAAGTTTATATTTAGTTTT 7103
: : : : : : : : : : : : : : : : : : : : : : : : : : :
302 rPheAlaThrLeuSerPheLeuAlaGluThrPheIlePhe.....L 316
: : : : : : : : : : : : : : : : : : : : : : : : : : :
7104 ATTGCTGGCTTATTCCTAATTAATCAATCATCGCGTTTTTAAATGTCAT 7153
: : : : : : : : : : : : : : : : : : : : : : : : : : :
316 euTyrValGlyMetAspAlaLeuAsp..... 324
: : : : : : : : : : : : : : : : : : : : : : : : : : :
7154 TATTTATGGCGTGTGATCATGATGATCAGGATCGGACACGCGATATT 7203
: : : : : : : : : : : : : : : : : : : : : : : : : : :
325 .....IleAspLysTrpArgSerValSerAspThrProGlyThrSerIl 339
: : : : : : : : : : : : : : : : : : : : : : : : : : :
7204 CGTTTGCTAAATGGTATGCTGAAGCTCC.....CCTAAAT 7241
: : : : : : : : : : : : : : : : : : : : : : : : : : :
339 eaValSerSerIleLeuMetGlyLeuValMetValGly.....A 353
: : : : : : : : : : : : : : : : : : : : : : : : : : :
7242 GCATATTGTCATGTGCTGCTGCAATATCTATGCGAGGTGTCCTTTT 7291
: : : : : : : : : : : : : : : : : : : : : : : : : : :
353 rgAlaAlaPheValPheProLeuSerPheLeuSerAsnLeuAlaLysLys 369
: : : : : : : : : : : : : : : : : : : : : : : : : : :
7292 TAAATGGCTTTTAAAGTAGGAATGTTTTAGATTCGTAACTAAGCA 7341
: : : : : : : : : : : : : : : : : : : : : : : : : : :
370 AsnGlnSerGluLysIleAsnPheAsnMetGlnValIleThrTrpSe 386
: : : : : : : : : : : : : : : : : : : : : : : : : : :
7342 AACGAACCTTGATCAATATGCTTCGTATTAACGTTTGTGATTTTCAAT 7391
: : : : : : : : : : : : : : : : : : : : : : : : : : :
386 rGlyLeuMetArgGlyAlaValSerMet.....AlaLeuAlaTyrAsnL 401
: : : : : : : : : : : : : : : : : : : : : : : : : : :
7392 AGGTGTCATCGCGAGTATATTGACTTTTACTTATGCACCTTTACATGATA 7441
: : : : : : : : : : : : : : : : : : : : : : : : : : :
401 ysPheThrArgAlaGlyHisThrAspValArgGly..... 412
: : : : : : : : : : : : : : : : : : : : : : : : : : :
7442 AAGAACATTCGGGGAATTTACATATAGAAAAATTTAAACGTAACAA 7491
: : : : : : : : : : : : : : : : : : : : : : : : : : :
413 .....AsnAlaIleMetIleThrSerTh 420
: : : : : : : : : : : : : : : : : : : : : : : : : : :
7492 ATACATGAACCATGGCTATTTAGTTTACCAGCTGTGATTTAATGTTACT 7541
: : : : : : : : : : : : : : : : : : : : : : : : : : :
420 rIleThrValCysLeuPheSerThrValValPheGly 432
: : : : : : : : : : : : : : : : : : : : : : : : : : :
7542 CATTCAGTATTCTCTTTGTTCCAAACGTTTTTGGC 7578
: : : : : : : : : : : : : : : : : : : : : : : : : : :
seq_name: /cgn2_6/ptodata/1/ina/5A_COMB.seq:US-08-920-812-6
seq_documentation_block:
; Sequence 6, Application US/08920812
; Patent No. 5763188
; GENERAL INFORMATION:
; APPLICANT: Ohno, Tsuneya
; APPLICANT: Matsuhisa, Akio
; APPLICANT: Uehara, Hirotugu
; APPLICANT: Eda, Soji
; TITLE OF INVENTION: Probe for Diagnosing Infectious Disease
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
```

```
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/920,812
; FILING DATE: 29-AUG-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/362,577
; FILING DATE: 27-MAR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Rin-Laures, Li-Hsien
; REGISTRATION NUMBER: 33,547
; REFERENCE/DOCKET NUMBER: 19036/32420
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312/474-6300
; TELEFAX: 312/474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8654 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: Genomic DNA
; ORIGINAL SOURCE:
; ORGANISM: Staphylococcus epidermidis
; STRAIN: Clinical isolate SE-22
; US-08-920-812-6

alignment_scores:
    Quality: 122.00      Length: 695
    Ratio: 0.391        Gaps: 36
    Percent Similarity: 44.892    Percent Identity: 18.849

alignment_block:
US-09-271-584A-2 x US-08-920-812-6/rev ..

Align seg 1/1 to reverse of: US-08-920-812-6 from: 1 to: 8654

14 ThrSerAspHisAlaSerValValAlaLeuAsnLeuPheValAlaLeuLe 30
||||| : : : : : : : : : : : : : : : : : : : : : : : : :
4665 ACCTCGGAACATTCACAACTT.....ATTCTATACT 4634
||||| : : : : : : : : : : : : : : : : : : : : : : : : :
30 ucysAlaCysIleValLeuGlyHisLeuLeuGluGluAsnArgTrpMetA 47
||||| : : : : : : : : : : : : : : : : : : : : : : : : :
4633 TTGCTTA.....TGGCTCA 4620
||||| : : : : : : : : : : : : : : : : : : : : : : : : :
47 snGluSerIleThrAlaLeuLeuIleGlyLeuGlyThrGlyValThrIle 63
||||| : : : : : : : : : : : : : : : : : : : : : : : : :
4619 ATCATTTTAGAAACATGATTGATGATG..... 4588
||||| : : : : : : : : : : : : : : : : : : : : : : : : :
64 LeuLeuIleSerLysGlyLysSerHisLeuLeuValPheSerGluAs 80
||||| : : : : : : : : : : : : : : : : : : : : : : : : :
4587 ATAATCATATTAAATAGAAAAGTTGGTTTTTA.....TC 4553
||||| : : : : : : : : : : : : : : : : : : : : : : : : :
80 pleuPhePheIleTyrLeuLeuProIleIle.....P 92
||||| : : : : : : : : : : : : : : : : : : : : : : : : :
4552 ATTGTTTTTCTATATCTTTTAGCAATTTTGTGGCATGTTACTTCCTA 4503
||||| : : : : : : : : : : : : : : : : : : : : : : : : :
92 heAsnAlaGlyPheGlnValLysLysGlnPhePheArg...AsnPhe 107
||||| : : : : : : : : : : : : : : : : : : : : : : : : :
4502 TTCAAACTCAATCAACTCAAGGCTAAGCAATTTACCGTTTCATCTTT 4453
||||| : : : : : : : : : : : : : : : : : : : : : : : : :
```

seq\_name: /cgn2\_6/ptodata/1/ina/5A\_COMB.seq:US-08-920-827-6

```

seq_documentation_block:
; Sequence 6, Application US/08920827
; Patent No. 5770375
; GENERAL INFORMATION:
; APPLICANT: Ohno, Tsuneya
; APPLICANT: Matsuhisa, Akio
; APPLICANT: Uehara, Hirotugu
; APPLICANT: Eda, Soji
; TITLE OF INVENTION: Probe for Diagnosing Infectious Disease
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/920,827
; FILING DATE: 29-AUG-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/362,577
; FILING DATE: 27-MAR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Rin-Laures, Li-Hsien
; REGISTRATION NUMBER: 33,547
; REFERENCE/DOCKET NUMBER: 19036/32420
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312/474-6300
; TELEFAX: 312/474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8654 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: Genomic DNA
; ORIGINAL SOURCE:
; ORGANISM: Staphylococcus epidermidis
; STRAIN: Clinical Isolate SE-22
; US-08-920-827-6

alignment_scores:
Quality: 122.00 Length: 695
Ratio: 0.391 Gaps: 36
Percent Similarity: 44.892 Percent Identity: 18.849

alignment_block:
US-09-271-584a-2 x US-08-920-827-6/rev ..
Align seg 1/1 to reverse of: US-08-920-827-6 from: 1 to: 8654
14 ThrSerAspHisAlaSerValValAlaLeuAsnLeuPheValAlaLeuLe 30
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
4665 ACCTCGGAACATTCACAACTT.....ATTCTATATCT 4634
30 uCysAlaCysIleValLeuGlyHisLeuLeuGluGluAsnArgTrpMetA 47
|||||
4633 TTGCTTA.....TGCTCA 4620
47 snGluSerIleThrAlaLeuLeuIleGlyLeuGlyThrGlyValThrIle 63
|||||
4619 ATCATTTTGAACAACATGATTGAATGATG..... 4588
64 LeuLeuIleSerIlySerHisLeuLeuValPheSerGluAs 80

```

```

|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
4587 ATAATCATATTAAATGAGAAAGTTGGTTTTTA.....TC 4553
80 pLeuPhePheIleTyrLeuLeuProPheIle.....P 92
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
4552 ATTGTTTTTCTATATCTTTTAGGAATTTTGGTGGCATGTTACTTCCTA 4503
92 heAsnAlaGlyPheGlnValLysLysGlnPhePheArg....AsnPhe 107
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
4502 TTCAAACTTCAATCAACTCAAGCTTAAGCCAAATTTACCCGTTTCATCTTT 4453
108 ValThrIleMetLeuPheGlyAlaValGlyThrIleIleSerCysThrI 124
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
4452 TATGCTTCAACCATATCTTTTGGTGTGCGAACATA.....TCCTTACT 4409
124 eile.....SerLeuGlyValThrGlnPhePheL 134
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
4408 TGTGCTTAATATATATTATTCACCAAGTATTAAACACCAAGATCTTTT 4359
134 yLysLeuAspIleGly...ThrPheAspLeuGlyAspTyrLeuAlaIle 149
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
4358 CTAAGCAACACACTTAACATATATCTTGGGTCTAGGAGGACTATTAGGTGT 4309
150 .....GlyAlaIlePh 153
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
4308 ATCTATTTAACCTGGAACCTTATTATTATTCACCAAGATTAGCGCGACCT 4259
153 eAlaAlaThrAspSerValCysThrLeuGlnValLeuAsnGln..... 167
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
4258 TACTGTTGTTATTACGGTT...ACAGGCAAAATTTATTATGGGTGTAATCA 4212
168 ..AspGluThrProLeuLeu.....TyrSerLeuVal 177
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
4211 TTGATACATTTTGGATTATTGGCGCCCATCAACAATCTTTTACATTATTT 4162
178 PheGlyGluGlyValValAsnAspAlaThrSerValValValPheAsnAl 194
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
4161 AAGGTGTTGGAATTTATTTTAACTAGGAATTTATTTTGAATTA 4112
194 aileGlnSerPheAspLeuThr...HisLeuAsnHisGluAlaAlaPheH 210
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
4111 TGTTCGTAGACATCTCTGTTAATAGACATAAGAATACACCAATAGTATCT 4062
210 isLeuLeuGlyAsnPheLeuTyrLeuPhe..... 219
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
4061 GGTGCTTATTGGATTCTGTTATTTGG..TTTCGTCCGCCAATCCAAACA 4013
220 .....LeuLeuSerThrLe 224
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
4012 ATCAACAGCACACTGGCTCAACACACACATCATCTATTTTGGTTCACCT 3963
224 uLeuGlyAlaAlaThrGlyLeuIleSerAlaTyrValIleLysLysLeu 241
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
3962 CATATCATTTAGCGTGTGTACAATAGCGTTATTATTATTA...ACATTAG 3916
241 yrPheGlyArg.....HisSerThrAspArgGluValAlaLeu 253
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
3915 TATTTAATCGTAGTCTGAAAAATATCATCTACTCATAAAACACTAGGTAAG 3866
254 MetMetLeuMetAlaTyrLeuSerTyrMetLeuAlaGluLeuPheAspLe 270
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
3865 ATTAATCTATTATTATTATTTAGGAGGACTACTAGGAATGGCATTTCGTAAC 3816
270 uSerGlyIleLeuThrValPhePhe..... 278
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
3815 ATCTAATATTATTATTATGCTTCTTAGGAGCAGCTTTAACTACTATAA 3766
279 .....CysGlyIleValMetSerHisTyrThr 287
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
3765 TTGCTATGATGGCCAAATGATTATGGGTATTATTATTCGATCATTTCCGGT 3716
288 TrpHisAsnValThrGluSerSerArgIleThrThrLysHisThrPheAl 304
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```



[illegible]

```
476 .....SerGlyAsnHisAsnValProArgProAspSerIleAr 488
|||||.....|
2981 CAAATAACACTTAGTAATGAACGTAACAACAAATTTTAAATCT..... 2937
488 gGlyPheLeuThrArgThrValHisThrValHisTyrTyrArgGlnP 505
2936 .....AAAAATGATTACATATTAT..... 2916
505 heAspAspSerPheMetArg..... 511
|||||.....|
2915 .....TCATTCCTTAATCAGCTAAATTAACAATGATTCACATAA 2874
512 .....ProValPheGlyGlyArgGlyP 519
|||||.....|
2873 TCCTCAATTAATCCACTTTATTTTACCAATTAATGGAAGTATTCATA 2824
519 evalProPheValProGlySerProThrGluArg 530
2823 TATTAAATCTATCCCTTCAAAACCAACGAAAG 2790
seq_name: /cgn_6/ptodata/1/ina/5A_COMB.seq:US-08-362-577C-6
```

```
seq_documentation_block:
; Sequence 6, Application US/08362577C
; Patent No. 5807673
; GENERAL INFORMATION:
; APPLICANT: Ohno, Tsuneya
; APPLICANT: Matsuhisa, Akio
; APPLICANT: Uehara, Hirotugu
; APPLICANT: Eda, Soji
; TITLE OF INVENTION: Probe for Diagnosing Infectious Disease
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSER: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/362.577C
; FILING DATE: 27-MAR-1995
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Rin-Laures, Li-Hsien
; REGISTRATION NUMBER: 33,547
; REFERENCE/DOCKET NUMBER: 19036/32420
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312/474-6300
; TELEFAX: 312/474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8654 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: Genomic DNA
; ORIGINAL SOURCE:
; ORGANISM: Staphylococcus epidermidis
; STRAIN: Clinical Isolate SE-22
US-08-362-577C-6
```

```
alignment_scores:
  Quality: 122.00      Length: 695
  Ratio: 0.391        Gaps: 36
Percent Similarity: 44.892 Percent Identity: 18.849
```

```
alignment_block:
US-09-271-584A-2 x US-08-362-577C-6/rev ..
Align seg 1/1 to reverse of: US-08-362-577C-6 from: 1 to: 8654

14 ThrSerAspHisAlaSerValValAlaLeuAsnLeuPheValAlaLeuLe 30
|||||.....|
4665 ACCTCGGACATTCACACTT.....ATTTCTATACT 4634
30 uCysAlaCysIleValLeuGlyHisLeuGluGluAsnArgTrpMetA 47
|||||.....|
4633 TTGCTTA.....TGCTCA 4620
47 snGluSerIleThrAlaLeuLeuGlyLeuGlyThrGlyValThrIle 63
|||||.....|
4619 ATCATTTTAGAACATGATTTATGAATGCTATG..... 4588
64 LeuLeuIleSerLysGlyLysSerHisLeuLeuValPheSerGluAs 80
|||||.....|
4587 ATAATCATATTAAATGAGAAAGTTGGTTTAA.....TC 4553
80 pleuPhePheIleThrLeuLeuProIleIle.....P 92
|||||.....|
4552 ATTTGTTTTTCTATATCTTTTAGGAATTTTGTGGCATGTTACTTCTCA 4503
92 heAsnAlaGlyPheGlnValLysLysLysGlnPhePheArg....AsnPhe 107
|||||.....|
4502 TTCAAACTTCAATCAACTCAAGGCTAAGCCAATTTTACCCGTTTCATCTTT 4453
108 ValThrIleMetLeuPheGlyValGlyThrIleIleSerCysThrI 124
|||||.....|
4452 TATGCTTCAACCATATCTTTTGCCTCGGAACAATA.....TGCCTACT 4409
124 eIle.....SerLeuGlyValThrGlnPhePheL 134
|||||.....|
4408 TGTGCTTAATATCATTTATTCATCCACAGTATTAAACACCAGATTTCTTT 4359
134 ysLysLeuAspIleGly....ThrPheAspLeuGlyAspTyrLeuAla 149
|||||.....|
4358 CTAGCAAAACACTTAACTATACTTGGGTCTTAGGAGGACTATTAGGTGT 4309
150 .....GlyAlaIlePhe 153
|||||.....|
4308 ATCTATTTTAACTGGAACTTATTATTATTACCAAGATTAGCGCAGCACT 4259
153 eAlaAlaThrAspSerValCysThrLeuGlnValLeuAsnGln..... 167
|||||.....|
4258 TACTGTTGTTATTACGGTT...ACAGGCAAAATTTATTATGGGTGTAATCA 4212
168 ..AspGluThrProLeuLeu.....TyrSerLeuVal 177
|||||.....|
4211 TTGATACATTTGGATTATTGGGCCCCCATCAACAATCTTTTACATTATTT 4162
178 PheGlyGluGlyValValAsnAspAlaThrSerValValValPheAsnAl 194
|||||.....|
4161 AAAGGTGTGGAATTATATTTTAACTACTGGAATTTATATTATGAATTA 4112
194 alleGlnSerPheAspLeuThr...HisLeuAsnHisGluAlaAlaPheH 210
|||||.....|
4111 TGTTCTGATAGACATCTCTTAAATAGACATAAGAATACACCAATAGTATCT 4062
210 isLeuLeuGlyAsnPheLeuTyrLeuPhe..... 219
|||||.....|
4061 GGTTCGTTATTGGATTGTTATTTGG..TTTCGTCGCCCAATCCAAACA 4013
220 .....LeuLeuSerThrLe 224
|||||.....|
4012 ATCAACAGCACACTGGCTCAACACACACACTCATCCATATTTTGGCTTCACT 3963
224 uLeuGlyAlaAlaThrGlyLeuIleSerAlaTyrValIleLysLysLeuT 241
|||||.....|
3962 CATATCATTTAGCGTGGTCAATAGCGTTATTATATTA...ACATTAG 3916
```

```
241 yrPheGlyArg.....HisSerThrAspArgGluValAlaLeu 253
|||||
3915 TATTATTCGTAGTCTGAAATATATCATCTACTATAAAACACATAGGTAAG 3866
254 MetMetLeuMetAlaTyrLeuSerTyrMetLeuAlaGluLeuPheAspLe 270
:::
3865 ATTAAATCTATTATTATTATGGAGGATAGTGGATGGCATTCTGTAAC 3816
:::
270 userGlyLeuThrValPhePhe..... 278
|||||
3815 ATCTAATATTATTATTAAATGCTTCTTAGGAGCAGCTTTAACTACTATAA 3766
279 .....CysGlyIleValMetSerHisTyrThr 287
3765 TTCTGATGATGGCCCAAAATGATTATGGTATTATTATCATCATCTTCGGT 3716
288 TrpHisAsnValThrGluSerSerArgIleThrLysHisThrPheAl 304
:::
3715 TTCTTGGGT...TCACCTAAACAAATAACATCACGT...AAGTTAGG 3672
304 aThrLeuSerPheLeuAlaGluThrPheIlePheLeu..... 316
:::
3671 TGGATTATTATGATAGCTAGTATTATTATTATTACGATTATTTTTAA 3622
317 .....TyrValGlyMetAspAla... 322
3621 TTCTCTTATTCGTAATAAAATATTCGCTTATATTATTAATGATAGTCTCTCA 3572
323 .....LeuAspIleAspLys.....Tr 328
3571 TACAATAAATGCTATTTCGAACTGCCAACAAAGATTTCGTGTGTTG 3522
328 pArg...SerValSerAspThrProGlyThrSerIleAlaValSerSerI 344
:::
3521 GCAGTTTTTATGAGTGATTCTCTTATAGTACATTTATGTAGAAATCATA 3472
344 eLeuMetGlyLeuValMetValGlyArgAlaAlaPheValPhePro.... 359
3471 TATTTTAAATAA.....AAGATAGCCTCTTATTTTCAATTT 3434
360 .....LeuSerPheLeuSerAsnLeuAlaLysLysAsnGlnSerGluLys 374
3433 ATTAATGATTGATTACITTTATTTTGAATTAACAACATATACAAACCAA 3384
375 IleAsnPheAsnMetGlnValValIleTrp..... 384
3383 AGTAATGCTTATTAAAAATTAATATATTTCATTTAAAAATAATAGTGATA 3334
385 .....TrpSerClyLeuMetArgGlyAlaValSerMetAlaLeuAlat 399
3333 ATACCTCAACAATAATATTATGAGG..... 3309
399 yrAsnLysPheThrArgAlaGlyHisThrAspValArgGlyAsnAlaIle 415
3308 .....TGAACATCCATGACCAAGCAACGCCATC 3282
416 MetIleThrSer.....ThrIleThrValCysLeuPheSe 427
3281 GTTATTTCTTCATTACTATTATTTTCATTTCTCACAACCTTGATGTTTAG 3232
427 rThrValValPheGlyMetLeuThrLys..... 436
3231 TTACATTACTTTCTCTAAATTAATCTTCTAAACAGAGATATTAAAAATA 3182
436 ..... 436
3181 TAAAAACTATAAAGCAAAACCCAAATTAATCTCTTCAACAACCTGGTATA 3132
437 .....ProLeuIleSerTyrLeuLeu.....ProHisGlnAsnAl 448
3131 CAAGCCCTCTTAACAGCTATATTATTTCTACTACACTCCCACTTTTAGGACA 3082
```

```
448 aThrThrSerMetLeuSerAspAsnThrProLysSerIleHisIleP 465
3081 ACTCAAAACACTTTTAGTAAAAAATCATAGCAAGATATCTAAAAACACTC 3032
465 roLeuLeu.....AspGlnAspSerPheIleGluPro... 475
3031 CAGTGTAGATATTACAAATTATGAAAAAAGATCATTTATATGAGCAACTC 2982
476 .....SerGlyAsnHisAsnValProArgProAspSerIleAr 488
2981 CAAATAACACTTAGTAAATGAACGTAAACAAACAAACATTTAAATCT... 2937
488 gGlyPheLeuThrArgProThrArgThrValHisTyrTyrTrpArgGlnP 505
2936 .....AAAAATGATTTACATTTATTAT... 2916
505 heAspAspSerPheMetArg..... 511
2915 .....TCATTCCTTAATCAGCTAAATATACAAATGATTCAAACTAAA 2874
512 .....ProValPheGlyGlyArgGlyPh 519
2873 TCCTCAATTAATCCACTTTATTTTCACCAATTAATGGAAGTATTTCTATA 2824
519 eValProPheValProGlySerProThrGluArg 530
2823 TATTAAATCTATCCCTTCAAAACCAACGAAGAAG 2790
seq_name: /cgn2_6/ptodata/1/ina/5B_COMB.seq:US-08-920-828-6

seq documentation block:
; Sequence 6, Application US/08920828
; Patent No. 5853998
; GENERAL INFORMATION:
; APPLICANT: Ohno, Tsuneya
; APPLICANT: Matsuhisa, Akio
; APPLICANT: Uehara, Hirotsugu
; APPLICANT: Eda, Soji
; TITLE OF INVENTION: Probe for Diagnosing Infectious Disease
; NUMBER OF SEQUENCES: 25
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/920,828
; FILING DATE: 29-AUG-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/362,577
; FILING DATE: 27-MAR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Rin-Laures, Li-Hsien
; REGISTRATION NUMBER: 33,547
; REFERENCE/DOCKET NUMBER: 19036/32420
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312/474-6300
; TELEFAX: 312/474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8654 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
```

```

; MOLECULE TYPE: Genomic DNA
;
; ORGANISM: Staphylococcus epidermidis
; STRAIN: Clinical Isolate SE-22
; US-08-920-828-6

alignment_scores:
    Quality: 122.00      Length: 695
    Ratio: 0.391        Gaps: 36
    Percent Similarity: 44.892      Percent Identity: 18.849

alignment_block:
    US-09-271-584A-2 x US-08-920-828-6/rev ..

Align seg 1/1 to reverse of: US-08-920-828-6 from: 1 to: 8655

14 ThrSerAspHisAlaSerValValAlaLeuAsnLeuPheValAlaLeuLe 30
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
4665 ACCTCGAACATTCACAACCT.....ATTCTATATC 4634

30 uCysAlaCysIleValLeuGlyHisLeuLeuGluGluAsnArgTtpMetA 47
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
4633 TTGCTTA.....TGGCTCA 4620

47 snGluSerIleThrAlaLeuLeuIleGlyLeuGlyThrGlyValThrIle 63
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
4619 ATCATTTTGAACAACATGATTATGAATGGTATG.....4588

64 LeuLeuIleSerLysGlyLysSerHisLeuLeuValPheSerGluAs 80
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
4587 ATAATCATATTAATTTAGAAAGTTGGTTTTTA.....TC 4553

80 pLeuPhePheIleTyrLeuLeuProPheIle.....P 92
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
4552 ATGTGTTTTTCTATATCTTTTAGGAATTTTGTGGCATGTTACTTCCTA 4503

92 heAsnAlaGlyPheGlnValLysLysLysGlnPhePheArg...AsnPh 107
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
4502 TTCAAACTTCAATCAACCAAGCGCTAAGCAATTTACCGGTTTCATCTTT 4453

108 ValThrIleMetLeuPheGlyAlaValGlyThrIleIleSerCysThrI 124
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
4452 TATGCTTCAACCATATCTTTGCTGCTCGGAACATA.....TGCTTACT 4409

124 eile.....SerLeuGlyValThrGlnPhePheL 134
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
4408 TGTGCTTAATATCATATTATTCATCCACAAGATTAAACACCAAGATTCIT 4359

134 ysLysLeuAspIleGly...ThrPheAspLeuGlyAspTyrLeuAlaIle 149
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
4358 CTAAGCAAAACACTTAACATATACTTGGGTCCTAGGAGGACTATTAGGTGT 4309

150 .....GlyAlaIlePh 153
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
4308 ATCTATTTAACTGGAACTTATTATTACCAAGATTAGGCGCAGACT 4259

153 eAlaAlaThrAspSerValCysThrLeuGlnValLeuAsnGln..... 167
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
4258 TACTGTTCTTATTACGGTT...ACAGGCAAAATTTATTGGGTGTAATCA 4212

168 ..AspGluThrProLeuLeu.....TyrSerLeuVal 177
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
4211 TTGATACATTTGGATATTGGCGGCCCATCAACAATCTTTTACATATTAT 4162

178 PheGlyGluGlyValValAsnAspAlaThrSerValValValPheAsnAl 194
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
4161 AAAGTGCTGGNAATTATATTTTAAATTACTGGAATFATATTATGAAATTA 4112

194 alleGlnSerPheAspLeuThr...HisLeuAsnHisGluAlaAlaPheH 210
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
4111 TGTTCGTAGACATCTGTTAATAGACATAAGAAATACACCAATAGTATCTT 4062

```





```

1156 .....CACTGATG.....TGTATATATT..... 1176
224 uLeuGlyAlaAlaThrGlyLeuIleSerAla.....TyrValIleL 238
1177 .CTGGCTGGGATTCCATTGATTCCTGCTTATACCGCTGTGGCA 1225
238 ystysLeuTyrPheGlyArgHis.....SerThrAspArgGluVal 251
1226 GAAGATTATATTACAATGACAACACTGCTGGATCAGTTCTGATACGCAACTC 1275
252 AlaLeuMetMetLeuMetAlaTyrLeuSerTyrMetLeuAlaGluLeu 268
1276 CTCTACATTATCCATGGCCCAATTTGTGCTTTATTGTGTAATCTTT 1325
268 eAspLeuSerGlyLeuLeuThrValPhePheCysGlyIleValMetSer 285
1326 TTTCCTATTAAATATTAGTGTCTTATCACCAGTTAAAGTTACT 1375
285 istYrThrTrpHisAsnValThrGluSerSerArgIleThrThrLysHis 301
1376 ACCAAGCAGAATCCAATCTGATACATGAAAGCT.....GTGAGAGCT 1416
302 ThrPheAlaThrLeuSerPheLeuAlaGluThrPheIlePheLeuTyrVa 318
1417 ACGCTTATCTGCTGGCGTGTGCTTGGCATTGAATTTGCTGATT..... 1461
318 lGlyMetAspAlaLeuAspIleAspLysTrpArgSerValSerAspThr 335
1462 .....CCATGGCGA.....C 1471
335 roGlyThrSerIleAla.....ValSerSerIleLeu 345
1472 CTGAAGGAAAGATTGCGAGAGAGATATGATTACATCAATAACATCTC 1521
346 Met.....GlyLeuValMetValGlyArgAlaAlaPheValPhePr 359
1522 ATGCATTATCCTGCTGGCGTGTGCTTGGTA.....TCTACAAATTTCTGCTCTT 1565
359 oLeuSerPheLeuSerAsnLeuAlaLysLysAsnGlnSerGlu...LysI 375
1566 TAATGGAGAGGTTCAAGCAATTCGAGAGAAACTGGAATCAATATAMA 1615
375 leAsnPheAsnMetGlnValIleTrpTrpSerGlyLeuMetArgGly 391
1616 TCCAATTTGGA..... 1626
392 AlaValSerMetAlaLeuAlaTyrAsnLysPheThrArgAlaGlyHisTh 408
1627 .....ACAACCTTTCC.....CACTC 1643
408 rAspValArgGlyAsnAlaIleMetIleThrSerThrIleThr 422
1644 AGATACCTCGTAGTGCACTTACACGGTTTCAACAATCAGC 1686

seq_name: /cqn2_6/ptodata/1/ina/5A_COMB.seq:us-08-098-141-1
seq_documentation_block:
; Sequence 1. Application US/08098141
; Patent No. 5441875
; GENERAL INFORMATION:
; APPLICANT: Hediger Ph.D., Matthias A.
; TITLE OF INVENTION: Urea Transporter Polypeptide
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Choate, Hall & Stewart
; STREET: Exchange Place, 53 State Street
; CITY: Boston
; STATE: MA
; COUNTRY: U.S.A.
; ZIP: 02119
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk

```

```

; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/098,141
; FILING DATE: 19930723
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Kaplan Esq., Warren A.
; REGISTRATION NUMBER: 34,199
; REFERENCE/DOCKET NUMBER: 092662-001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617:227-5020
; TELEFAX: 617:227-7566
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3060 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHEetical: NO
; ANTI-SENSE: NO
; US-08-098-141-1

alignment_scores:
Quality: 98.50 Length: 319
Ratio: 0.724 Gaps: 17
Percent Similarity: 42.633 Percent Identity: 21.317

alignment_block:
US-09-271-584A-2 x US-08-098-141-1 ..
Align seg 1/1 to: US-08-098-141-1 from: 1 to: 3060

34 lleValLeuGlyHisLeuLeuGluGluAsnArgTrpMetAsnGluSerIle 50
|||||:||||| |||: ||| ||| |||
1262 ATCGTGATTGGC...CTCTTTGCCAAGCCATGTTGG.....GCCAT 1302

50 eThrAlaLeuLeuIleGlyLeuGlyThrGlyValThrIleLeuLeuIle 67
|||||:||||| |||: ||| ||| |||: |||
1303 CCGGGATGCTTGGCCACCGCATGTCCACCTTGACTGCCCTCATCTCGA 1352

67 eLysGlyLysSer.....His..... 73
|||: |||: |||
1353 GCCAGACAGGTACGCATCGCTCCGACTCCAGGCTACACGGCGTG 1402

74 .....LeuLeuValPheSerGlu.....AspLeuPhePh 83
|||: |||: |||: |||: |||: |||
1403 CTGGTGGGGTGTGATAGCGGTGTTCTCCGACAGGGCGACTATTACTG 1452

83 eileTyrLeuLeuProProlIleIlePheAsnAlaGlyPheGlnValLysL 100
: |||: |||: |||: |||: |||: |||
1453 GTGGTCTCTGCTCCCTGTCATCGTATGTC..... 1483

100 ysLysGlnPhePheArgAsnPheValThrIleMetLeuPheGlyAlaVal 116
1483 ..... 1483

117 GlyThrIleIleSerCysThrIleIleSerLeuGlyValThrGlnPhePh 133
: |||: |||: |||: |||: |||: |||
1484 .....ATGTCTGCCCATCTCTCCAGTGCCCTGGCACCATTCTT 1524

133 eLysLysLeuAspIleGlyThrPheAspLeuGlyAspTyrLeuAlaIleG 150
|||||: |||: |||: |||: |||: |||
1525 CAGCAAGTGGGACCTCCCGACTTCCACCTGCCCTCAACATCGCCGTGA 1574

150 lYalaIlePheAlaAlaThrAspSerValCysThrLeuGlnValLeuAsn 166
: |||: |||: |||: |||: |||: |||
1575 CCCTGTACCTGGCACCACAGGCCAC..... 1600

167 GlnAspGluThrProLeuLeuTyrSerLeuValPheGlyGluValVa 183

```

```
1601 .....TACAACTCTTCCTCCCAACACAGTCT 1629
183 lasnaspAlaThrSerVal.....ValValPheAsnAlaIleGlnSerP 198
1630 GCAGCCTGTATCCTCGGCGCAACATCACCTGTGTGAGACAG..... 1675
198 heAspLeuThrHisLeuAsnHisGluAlaAlaPheHisLeu..... 212
1676 .....GTGCCCTTCCTCTTGAGA 1693
212 ..... 212
1694 GCATCCCGCTTGGAACTGGCAAGTGTATGCTGTGTGATAACCCCTGGAC 1743
213 ....GlyAsnPheLeuTyr.....LeuPheLeuLeuSerThrLeu..... 224
1744 TGGAGGCATTTTCCTCATCGCTCTCTTTCATATCTTCACCTCTTATTGCT 1793
225 .....LeuGlyAlaAlaThrGlyLeuIleSerAlaTyrValIle 237
1794 TGCAATCGGCAATTTGGATCCACATGGGGATGTTAGCAGACACTCACGATT 1843
238 .....LysLysLeuTyrPheGlyArgHisSerThrAspArgL 250
1844 GCCACACCTTTGACTCCATCTACTTGGCCTGTGTGCTTCAACAGCAC 1893
250 uValAla.....LeuMetMetLeuMetAlaTyrLeuS 261
1894 GCTGGCATCATCGCTGTGGAGGAGCATTTCTACGTCATCACCTGGCAGA 1943
261 erTyrMetLeuAla.....GluLeuPhe..... 268
1944 CTCACCTCTTCTGTGCTGCTGCGCCCTGTGTGGGCTACGTGGTGCT 1993
269 AspLeuSerGlyIleLeuThrValPhe..... 277
1994 GCCTGACCAAGCATATTATCTGTGTGTGGATTACCAACCTGCACCTGGCC 2043
278 .PheCys 279
2044 CTCTCTGC 2050
```

seq\_name: /cgn2\_6/ptodata/1/ina/6\_COMB.seq:US-08-602-809-3

seq\_documentation\_block:

```
; Sequence 3, Application US/08602809
; Patent No. 6008012
; GENERAL INFORMATION:
; APPLICANT: BERGSMÄ, DERK
; APPLICANT: ELLIS, CATHERINE
; TITLE OF INVENTION: HUMAN SOMATOSTATIN-LIKE R
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Ratner & Prestia
; STREET: P.O. Box 980
; CITY: Valley Forge
; STATE: PA
; COUNTRY: US
; ZIP: 19482-0980
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/602,809
; FILING DATE: 13-JUN-1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCW/US95/16472
; FILING DATE: 15-DEC-1995
```

```
; APPLICATION NUMBER: US 08/357,675
; FILING DATE: 16-DEC-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Prestia, Paul F
; REGISTRATION NUMBER: 23,031
; REFERENCE/DOCKET NUMBER: P50277
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 601-407-0700
; TELEFAX: 610-407-0701
; TELEX: 846169
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1080 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; US-08-602-809-3

alignment_scores:
  Quality: 96.50      Length: 371
  Ratio: 0.652       Gaps: 22
Percent Similarity: 39.892 Percent Identity: 22.642

alignment_block:
US-09-271-584A-2 x US-08-602-809-3 ..
Align seg 1/1 to: US-08-602-809-3 from: 1 to: 1080

171 ProLeuLeuTyrSerLeuValPheGlyGluGlyValValAsnAspAlaTh 187
    ||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
94  CCTACCTCTACAGCATCATCTTGTGTGGTGGGAATA..... 129

187 rSerValValValPheAsnAlaIleGlnSerPheAspLeuThrHisLeuA 204
129 ..... 129

204 snHisGluAlaAlaPheHisLeuLeuGlyAsnPheLeuTyrLeuPheLeu 220
    ||| ||||| ||||| ||||| ||||| ||||| |||||
130 .....TTGGAAACAGCTTGTGTGGTGCATCTCTC 156

221 LeuSerThrLeuLeuGlyAlaAlaThrGlyLeuIleSerAlaTyrValIl 237
    ||| ||||| ||||| ||||| ||||| ||||| |||||
157 ATT.....TACTTTTA 167

237 elysLysLeuTyrPheGlyArgHisSerThrAspArgGluValAlaLeuM 254
    ||||| ||||| ||||| ||||| ||||| ||||| |||||
168 CATGAAGCTG.....AAGACTGTGCCACGC 193

254 etMetLeuMetAlaTyrLeuSerTyrMetLeuAlaGluLeuPheAspLeu 270
    ||||| ||||| ||||| ||||| ||||| ||||| |||||
194 TCTTTCTT.....CTCAATCTCGCTTGGCT.....GACTTA 225

271 SerGlyIleLeuThrValPhePheCysGlyIleValMetSer...HisTy 286
    ||||| ||||| ||||| ||||| ||||| ||||| |||||
226 TGTCTTTTGTGACTTGTGCCCTGTGGCAGCTCTATACCGCTATGAGTA 275

286 rThrThrHisAsnValThrGluSerSerArgIleThrThrLysHisThrP 303
    ||| ||||| ||||| ||||| ||||| ||||| |||||
276 CCGCTGGCCCTTCGGCAATCACCTATGTAGATCGCTTCG..... 315

303 heAlaThrLeuSerPheLeuAlaGluThrPheIlePheLeuTyrValGly 319
    ||||| ||||| ||||| ||||| ||||| ||||| |||||
316 ..GCCAGCTGAGCTTCAACCTCTACGCCAGCTGTGTGCTCTT..... 354

320 MetAspAlaLeuAspIleAspLysTrpArgSerValSerAspThrProGl 336
    ||||| ||||| ||||| ||||| ||||| ||||| |||||
355 CTCACGTGTCTCAGCATCGACCGCTAC..... 381

336 yThrSerIleAlaValSerSerIleLeuMetGlyLeuValMetValGlyA 353
381 ..... 381
```

```

353 rgAlaAlaPheValPheProLeuSerPheLeuSerAsnLeuAlaLysLys 369
|||:::|||||:::|||||:::
382 .CTGCCCATCTGCACCAATGAAG.....TCTGCCTTCGCCGACG 423
|||:::|||||:::|||||:::
370 AsnGlnSerGluLysIleAsnMetGlnValValIleThrPheSe 386
|||:::|||||:::|||||:::
424 ATCCTGGTGGCCAAAGTCACTGC.....ATCATCATCTGG..... 459
|||:::|||||:::|||||:::
386 rGlyLeuMetArgGlyAlaValSerMetAlaLeuAlaTyrAsnLysPheT 403
|||||:::|||||:::|||||:::
460 ....CTGATPCTGCTGCTGGCCATTTGCCAGCTGTTCATCCACCGA.... 501
|||:::|||||:::|||||:::
403 hrArgAlaGlyHisThrAspValArgGlyAsnAlaIleMetIleThrSer 419
|||:::|||||:::|||||:::
502 .....AATGTATACTTCATCGAGAAC 522
|||:::|||||:::|||||:::
420 Thr....IleThrValCysLeuPhe.....SerThrVa 429
|||:::|||||:::|||||:::
523 ACCAATATACAGATGTGGCGCTTTCATATTAGTCTCGGAATTCGACGCT 572
|||:::|||||:::|||||:::
429 lValPheGlyMet....LeuThrLysProLeuIleSerTyrLeuLeuProH 445
|||:::|||||:::|||||:::
573 CCCATPAGGCGTGGCCCTTACCAAGAATATTCTGGGCTTCTTGTCCTT 622
|||:::|||||:::|||||:::
445 isGlnAsnAlaThrThrSer...MetLeuSerAspAsnThrProLys 460
|||:::|||||:::|||||:::
623 TCCTTATCATCTCACACGATATACCTTATTATTTGGAAAGCTCTAAAGAG 672
|||:::|||||:::|||||:::
461 SerIleHisIleProLeuLeuAspGlnAspSerPheIleGluProSerGl 477
|||:::|||||:::|||||:::
673 GCATTGAATTT..... 684
|||:::|||||:::|||||:::
477 yAsnHisAsnValProArgProAspSerIle.ArgGlyPheLeuThrArg 493
|||:::|||||:::|||||:::
685 .CAAAAGAACAAACCAAGAACGATGACATCTTTAGGATAATTA..... 727
|||:::|||||:::|||||:::
494 ProThrArgThrValHisTyrTyrTrpArgGln....PheAspAspSerPh 509
|||:::|||||:::|||||:::
728 .....TGGCGATTCTGCTTTCTTCTCTCTT 753
|||:::|||||:::|||||:::
509 eMetArgProValPheGlyGlyArgGlyPheValProPheValProGlyS 526
|||:::|||||:::|||||:::
754 T.....CTGGGT 761
|||:::|||||:::|||||:::
526 erProThrGlu 529
|||:::|||||:::|||||:::
762 CCCCCACCAAA 772
|||:::|||||:::|||||:::

seq_name: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:PCT-US95-16472-3
seq_documentation block:
; Sequence 3, Application PC/TUS9516472
; GENERAL INFORMATION:
; APPLICANT: Bergsma, Derk J
; APPLICANT: Ellis, Catherine E
; TITLE OF INVENTION: Human Somatostatin Receptor
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corporation/Corporate
; ADDRESSEE: Intellectual Property
; STREET: P. O. Box 1539-UW2220
; CITY: King of Prussia
; STATE: Pennsylvania
; COUNTRY: USA
; ZIP: 19406-0939
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/16472

```

Mon Mar 26 11:04:17 2001

```
353 rgAlaAlaPheValPheProLeuSerPheLeuSerAsnLeuAlaLysLys 369
    |||:|||| |||:||||
382 ..CTGCCATCGTCCACCAATGAAG.....TCTCGCTTCGCCGACG 423
    |||:|||| |||:||||
370 AsnGlnSerGluLysIleAsnPheAsnMetGlnValValIleTrpTrpSe 386
    |||:|||| |||:||||
424 ATGCTGGTGGCCAAAGTCACCTGC.....ATCATCATCTGG..... 459
    |||:|||| |||:||||
386 rgLysLeuMetArgGlyAlaValSerMetAlaLeuAlaTyrAsnLysPheT 403
    |||:|||| |||:||||
460 ....CTGATGGCTGGCTGGCCAGTTTCCACGCTGTCATCCACCGA.... 501
    |||:|||| |||:||||
403 hrArgAlaGlyHisThrAspValArgGlyAsnAlaIleMetIleThrSer 419
    |||:|||| |||:||||
502 .....AATGTATACTTTCATCCAGAAC 522
    |||:|||| |||:||||
420 Thr...IleThrValCysLeuPhe.....SerThrVa 429
    ||| ||| ||| ||| ||| |||
523 ACCAATATCACAGTGTGGCGCTTCATTATGAGTCTCGGAATTCGACGT 572
    |||:|||| |||:||||
429 IValPheGlyMet...LeuThrLysProLeuIleSerTyrLeuLeuProH 445
    : : : : : ||| ||| ||| : : : : :
573 CCCATAGGGCTGGCCCTACCAAGAAATATTCTGGGCTTCTGTCCCTT 622
    |||:|||| |||:||||
445 IsGlnAsnAlaThrThrSer...MetLeuSerAspAsnThrProLys 460
    ||| ||| ||| |||
623 TCCTTATCATCTCTACCGCTATACCTTATTGGAAGCTCTAAGAGAG 672
    |||:|||| |||:||||
461 SerIleHisIleProLeuLeuAspGlnAspSerPheIleGluProSerGI 477
    : : : : : ||| ||| |||
673 GCTTATGAAATT..... 684
    |||:|||| |||:||||
477 yAsnHisAsnValProArgProAspSerIle.ArgGlyPheLeuThrArg 493
    : : : ||| ||| ||| ||| ||| |||
685 .CAAAAGAACAAACCAAGAACGATGACATCTTAGGATAATTA..... 727
    |||:|||| |||:||||
494 ProThrArgThrValHisTyrTyrTrpArgGln...PheAspAspSerPh 509
    ||| ||| ||| ||| ||| |||
728 .....TGCGGATGTGCTTTTCTCTCTCTTT 753
    |||:|||| |||:||||
509 eMetArgProValPheGlyArgGlyPheValProPheValProGlyLys 526
    | ||| ||| |||
754 T.....CCTGGGT 761
    |||:|||| |||:||||
526 erProThrGlu 529
    ||| ||| ||| ||| ||| |||
762 CCCCCACCAAA 772
```



OM of: US-09-271-584A-2 to: GenEmbl.\* out\_format : pfs  
Date: Mar 26, 2001 10:35 AM  
About: Results were produced by the GenCore software, version 4.5,  
Copyright (c) 1993-2000 CompuGen Ltd.

Command line parameters:  
-MODEL=frame+p2n\_model -DEV=xlp  
-O=/cgn2\_1/USPto\_spool/US09271584/runat\_15032001\_151301\_5671/app\_query.fasta\_1.604  
-DB=GenEmbl -QFMT=fastcap -SUFFIX=p2n.rge -GAPOP=12.000  
-GAPEXT=4.000 -MINMATCH=0.100 -LOOPEL=0.000 -LOOPEXT=0.000  
-CGAPOP=4.500 -CGAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500  
-FCGAPOP=6.000 -FCGAPEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500  
-DELOP=6.000 -DELEXT=7.000 -START=1 -MATRIX=blotsum62  
-TRANS=human40.cdi -LIST=45 -DCCALIGN=200 -THR\_SCORE=pct  
-THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pfs  
-NORM=ext -MINLEN=0 -MAXLEN=2000000000  
-USER=US09271584\_cgn1\_1\_5190 -NCPU=6 -ICPU=3 -LONGLOG -NO\_XLPPY  
-WAIT -THREADS=1

Search information block:  
Query: US-09-271-584A-2  
Query length: 538  
Database: GenEmbl.\*  
Database sequences: 1118133  
Database length: -1736092196  
Search time (sec): 2605.960000

score\_list:

| Sequence          | Strd Orig | zscore  | EScore   | Len    | Documentation                     |
|-------------------|-----------|---------|----------|--------|-----------------------------------|
| gb_p11:AF106324   | + 2755.00 | 3469.57 | 4.2e-185 | 1614   | ! AF106324 Arabidopsis thaliana   |
| gb_p11:AF056190   | + 2755.00 | 3469.54 | 4.3e-185 | 1619   | ! AF056190 Arabidopsis thaliana   |
| gb_p13:TW021B04   | - 2282.50 | 2833.28 | 1.2e-149 | 90019  | ! AP007271 Arabidopsis thaliana   |
| gb_p11:AB038492   | + 2168.50 | 2725.00 | 1.3e-143 | 2637   | ! AB038492 Atriplex gmelini AgNH  |
| gb_p11:AB033989   | + 2041.00 | 2565.95 | 9.1e-135 | 2218   | ! AB033989 Ipomoea nil Purple m   |
| gb_p11:AB021878   | + 2023.50 | 2543.46 | 1.6e-133 | 2313   | ! AB021878 Oryza sativa OsNH1 m   |
| gb_p11:AC009465   | + 1924.00 | 2380.82 | 1.9e-124 | 92234  | ! AC009465 Arabidopsis thaliana   |
| gb_p11:ATAC011623 | + 1492.50 | 1837.47 | 3.4e-94  | 86022  | ! AC011623 Arabidopsis thaliana   |
| gb_p11:AB033990   | + 1266.50 | 1756.92 | 1.1e-79  | 7545   | ! AB033990 Ipomoea nil Purple ge  |
| gb_p11:AB015479   | - 1118.00 | 1365.83 | 6.4e-68  | 80675  | ! AB015479 Arabidopsis thaliana   |
| gb_in2:AF142676   | + 615.50  | 767.44  | 1.4e-34  | 2405   | ! AF142676 Drosophila melanogast  |
| gb_p12:AF279671   | + 586.00  | 747.76  | 1.7e-33  | 421    | ! AF279671 Mesembryanthemum crysu |
| gb_p15:DB7743     | + 581.00  | 717.84  | 7.9e-32  | 4408   | ! DB7743 Human mRNA for KIAA0267  |
| gb_p13:AF030409   | + 581.00  | 717.75  | 8.0e-32  | 4452   | ! AF030409 Homo sapiens sodium/H  |
| gb_in2:AF199463   | + 568.00  | 706.74  | 3.3e-31  | 2604   | ! AF199463 Drosophila melanogast  |
| gb_p12:SCD9461    | + 564.00  | 667.46  | 5.1e-29  | 78500  | ! AF199463 Drosophila melanogast  |
| gb_p12:AF279670   | + 542.00  | 691.40  | 2.4e-30  | 459    | ! AF279670 Mesembryanthemum cryst |
| gb_h1g6:AC017375  | + 530.00  | 625.24  | 1.1e-26  | 73515  | ! AC017375 Drosophila melanogast  |
| gb_in1:AF003590   | - 530.00  | 611.05  | 7.0e-26  | 301639 | ! AB003590 Drosophila melanogast  |
| gb_h1g1:AC008318  | - 513.00  | 588.47  | 1.3e-24  | 338007 | ! AC008318 Drosophila melanogast  |
| gb_p13:SPAC15A10  | + 502.50  | 595.79  | 5.0e-25  | 43702  | ! 297208 S.pombe chromosome I c   |
| gb_h1g6:AC016779  | - 486.00  | 563.30  | 3.2e-23  | 139626 | ! AC016779 Oryza sativa clone     |
| gb_om:RABNAHEX    | + 475.00  | 588.92  | 1.2e-24  | 2748   | ! M87007 Oryctolagus cuniculus N  |
| gb_ro:RATNHHEXIII | + 467.00  | 572.51  | 9.9e-24  | 5153   | ! M85300 Rat sodium-hydrogen exc  |
| gb_p13:AF111173   | + 466.50  | 575.04  | 7.1e-24  | 3761   | ! AF111173 Homo sapiens sodium/H  |
| gb_ro:AF100172    | + 466.50  | 574.00  | 8.2e-24  | 4174   | ! AF100172 Rattus norvegicus sod  |
| gb_p11:AR035574   | + 462.00  | 578.23  | 4.9e-24  | 2574   | ! AR035574 Sequence 8 from pater  |
| gb_p17:HSU28043   | + 462.00  | 573.14  | 9.1e-24  | 2584   | ! U28043 Human plasma membrane N  |
| gb_om:OPONHE3R    | + 459.00  | 568.34  | 1.7e-23  | 2859   | ! L42522 Didelphis virginiana Na  |
| gb_ro:RATMATSTOM  | + 455.50  | 560.74  | 4.5e-23  | 3925   | ! L11236 Rattus rattus mRNA sequ  |
| gb_om:RABNH2A     | + 454.50  | 559.77  | 5.1e-23  | 3813   | ! L13733 Rabbit ethylisopropylam  |
| gb_ro:RATNH2Z     | + 454.50  | 559.44  | 5.3e-23  | 3941   | ! L11004 Rat sodium-hydrogen exc  |
| gb_p13:AF073299   | + 452.50  | 560.68  | 4.5e-23  | 2711   | ! AF073299 Homo sapiens Na+/H+ e  |
| gb_in2:AF235935   | + 451.00  | 555.90  | 8.3e-23  | 3612   | ! AF235935 Drosophila melanogast  |
| gb_p11:AB023156   | + 451.00  | 551.61  | 1.4e-22  | 5537   | ! AB023156 Homo sapiens mRNA for  |
| gb_p13:AB023156   | + 444.00  | 550.13  | 1.7e-22  | 2665   | ! S81591 SLCA2-Na+/H+ exchanger   |
| gb_p13:AF141350   | + 443.00  | 549.36  | 1.9e-22  | 2538   | ! AF141350 Homo sapiens sodium/H  |
| gb_p13:AF141351   | + 443.00  | 549.36  | 1.9e-22  | 2538   | ! AF141351 Homo sapiens sodium/H  |
| gb_p13:AF141352   | + 443.00  | 549.36  | 1.9e-22  | 2538   | ! AF141352 Homo sapiens sodium/H  |
| gb_p13:AF141353   | + 443.00  | 549.36  | 1.9e-22  | 2538   | ! AF141353 Homo sapiens sodium/H  |

|   |  |                           |         |             |                               |
|---|--|---------------------------|---------|-------------|-------------------------------|
| gb_p13:AF141354   | + 443.00   | 549.36                    | 1.9e-22 | 2538        | ! AF141354 Homo sapiens sodiu |
| gb_p13:AF141355   | + 443.00   | 549.36                    | 1.9e-22 | 2538        | ! AF141355 Homo sapiens sodiu |
| gb_p13:AF141356   | + 443.00   | 549.36                    | 1.9e-22 | 2538        | ! AF141356 Homo sapiens sodiu |
| gb_p13:AF141357   | + 443.00   | 549.36                    | 1.9e-22 | 2538        | ! AF141357 Homo sapiens sodiu |
| gb_p13:AF141358   | + 443.00   | 549.36                    | 1.9e-22 | 2538        | ! AF141358 Homo sapiens sodiu |
| seq_name: gb_p11:AF106324   |  |                           |         |             |                               |
| seq_documentation_block:  |  |                           |         |             |                               |
| LOCUS AF106324  | 1614 bp  | mRNA                      | PLN     | 03-MAR-1999 |                               |
| DEFINITION Arabidopsis thaliana sodium proton exchanger Nhx1 mRNA, partial cds. |  |                           |         |             |                               |
| ACCESSION AF106324  |  |                           |         |             |                               |
| VERSION AF106324.1  | GI:4324596   |                           |         |             |                               |
| KEYWORDS  |  |                           |         |             |                               |
| SOURCE  | thale cress.   |                           |         |             |                               |
| ORGANISM  | Arabidopsis thaliana   |                           |         |             |                               |
| REFERENCE   | Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsi. |                           |         |             |                               |
| AUTHORS   | 1 (bases 1 to 1614)  |                           |         |             |                               |
| TITLE   | Gaxiola,R.A., Rao,R., Sherman,A., Grisafi,P., Alper,S.L. and Fink,G.R.   |                           |         |             |                               |
| JOURNAL   | The Arabidopsis thaliana proton transporters, AtNhx1 and Avp1, can function in cation detoxification in yeast  |                           |         |             |                               |
| MEDLINE   | Proc. Natl. Acad. Sci. U.S.A. 96 (4), 1480-1485 (1999)   |                           |         |             |                               |
| REFERENCE   | 99145575   |                           |         |             |                               |
| AUTHORS   | 2 (bases 1 to 1614)  |                           |         |             |                               |
| TITLE   | Gaxiola,R.A., Rao,R., Sherman,A., Grisafi,P., Alper,S.L. and Fink,G.R.   |                           |         |             |                               |
| JOURNAL   | Direct Submission  |                           |         |             |                               |
| FEATURES  | Submitted (12-NOV-1998) Whitehead, Nine Cambridge Center, Cambridge, MA 02142, USA   |                           |         |             |                               |
| SOURCE  | Location/Qualifiers  |                           |         |             |                               |
|   | 1..1614  |                           |         |             |                               |
|   | /organism="Arabidopsis thaliana"   |                           |         |             |                               |
|   | /cultivar="Columbia"   |                           |         |             |                               |
|   | /db_xref="taxon:3702"  |                           |         |             |                               |
|   | 1..>1614   |                           |         |             |                               |
|   | /codon_start=1   |                           |         |             |                               |
|   | /product="sodium proton exchanger Nhx1"  |                           |         |             |                               |
|   | /protein_id="AA016946.1"   |                           |         |             |                               |
|   | /db_xref="GI:4324597"  |                           |         |             |                               |
|   | /translation="MDDSLVSKPLSLSTSDHASVVALNLFVALLCACIVLGLLENR   |                           |         |             |                               |
|   | WMNESITALLIGLTGVTILISKSSHLVSEDLFFIYLPPIIFNAGFOVKKKQ  |                           |         |             |                               |
|   | FFNFVITMLFGAVGTIIISLGTQFFKKLDIGTDFLDGYLAIGAFATDSVC   |                           |         |             |                               |
|   | TLQVNODETPGLISLVGEGVVDATSVVFNAIQSFDTLHNHAAFLHGLNFYL  |                           |         |             |                               |
|   | FLLSLTILGAATGLISAVIKKLYFGRSTOREVALMMLAYLSYMLAELFDLGLILY  |                           |         |             |                               |
|   | FFCGIVMSHTWNVTSSRTTKHTFATLSLAETFIIFYVGMALDIDKWRVSDDT   |                           |         |             |                               |
|   | PCPSIAVSSILMLGLVAVGRAFPVPLSFLSNLAKKQSEKINFMQOVYIWSGLMGA  |                           |         |             |                               |
|   | VSMALATNKRTRAGTIDVRGNALMITSTIVCLFSTVVFQMLTKPLISYLPHONATT   |                           |         |             |                               |
|   | SMLSDDNTPKSIHIPLDQDSFTPEPSGNHNVPDPSIRGELTRTPRTTVHYWRQFDD   |                           |         |             |                               |
|   | FMRPVFGRGFVFPVPGSPTEPNPPDLKA"  |                           |         |             |                               |
| BASE COUNT  | 388 a 346 c 358 g 522 t  |                           |         |             |                               |
| ORIGIN  |  |                           |         |             |                               |
| alignment_scores:   |  |                           |         |             |                               |
|   | Quality: 2755.00   | Length: 538               |         |             |                               |
|   | Ratio: 5.121   | Gaps: 0                   |         |             |                               |
|   | Percent Similarity: 100.000  | Percent Identity: 100.000 |         |             |                               |
| alignment_block:  |  |                           |         |             |                               |
|   | US-09-271-584A-2 x AF106324  |                           |         |             |                               |
|   | Align seg 1/1 to: AF106324 from: 1 to: 1614  |                           |         |             |                               |
|   | 1 MetLeuAspSerLeuValSerLysLeuProSerLeuSerThrSerAspHis 17   |                           |         |             |                               |
|   |  |                           |         |             |                               |
|   | 1 ATGTTGGATTCCTCTAGTCGAAACTCGCTTCGTTATCGACATCTGATCA 50   |                           |         |             |                               |
|   |  |                           |         |             |                               |
|   | 17 salaSerValValAlaLeuAsnLeuPheValAlaLeuLeuCysAlaCysI 34   |                           |         |             |                               |
|   |  |                           |         |             |                               |

51 .CGCTTCTGCTGGTGGTGAATCTCTTTGTTGCACCTCTTTGTGCTTTGTA 100  
34 leValLeuGlyHisLeuLeuGluGluAsnArgTrpMetAsnGluSerIle 50  
101 TTGTTCTTGGTCATCTTTTGAAGAGATAGATGGATGAAGATCCATC 150  
51 ThrAlaLeuLeuIleGlyLeuGlyThrGlyValThrIleLeuLeuIleSe 67  
151 ACCGCTTTGTTGATTGGCTAGGCACCTGGTGTACCATTTTGTGATTAG 200  
67 rLysGlyLysSerSerHisLeuLeuValPheSerGluAspLeuPheI 84  
201 TAAAGGAAAAGCTCGCATCTTCGCTTTTGTAGTGAAGATCTTTCTTCA 250  
84 leTyrLeuLeuProPheIlePheAsnAlaGlyPheGlnValLysLys 100  
251 TATATCTTTTCCACCCCATATATATCAATGCGAGGTTTCAAGTAAAG 300  
101 LysGlnPhePheArgAsnPheValThrIleMetLeuPheGlyAlaValGI 117  
301 AAGCAGTTTTTCCGCAATTTCTGCTGACTATTATGCTTTTGGTGTGG 350  
117 yThrIleIleSerCysThrIleIleSerLeuGlyValThrGlnPhePheL 134  
351 GACTATTATTTCTGCAATATCATATCTCTAGGTGTAACACAGTTCTTTA 400  
134 ysLysLeuAspIleGlyThrPheAspLeuGlyAspTyrLeuAlaIleGly 150  
401 AGAGTTGACATTTGGAACCTTTGACTGGTGGTGATATCTTGTCTATTGGT 450  
151 AlaIlePheAlaAlaThrAspSerValCysThrLeuGlnValLeuAsnGI 167  
451 GCCATATTTTGTGCAACAGATTGATGTACTGCGAGGTTCTGAATCA 500  
167 nAspGluThrProLeuLeuTyrSerLeuValPheGlyGluGlyValVala 184  
501 AGACGAGACCTTTGCTTTTACACTTGTATTCGGAGAGGTTTGTGA 550  
184 snAspAlaThrSerValValValPheAsnAlaIleGlnSerPheAspLeu 200  
551 ATGATGCAACGCTCAGTTGTGCTTTCAACGCGATTGAGCTTTGATCTC 600  
201 ThrHisLeuAsnHisGluAlaAlaPheHisLeuLeuGlyAsnPheLeuTy 217  
601 ACTCACCTAAACACGAAAGCTGCTTTTTCATCTTCTTGAAACTTCTTGA 650  
217 rLeuPheLeuLeuSerThrLeuLeuGlyAlaAlaThrGlyLeuIleSerA 234  
651 TTTGTTTCTCCTAAGTACCTTGTGCTGGTGGTGCACACCGGCTGATAAGT 700  
234 laTyrValIleLysLysLeuTyrPheGlyArgHisSerThrAspArgGlu 250  
701 CGTATGTTATCAAGAGCTATATCTTTGGAAGGCACTCAACTGACCGAG 750  
251 ValAlaLeuMetMetLeuMetAlaTyrLeuSerTyrMetLeuAlaGluLe 267  
751 GTTGGCCCTTATGCTTATGGCGTATCTTCTTATATGCTTGTGCTGAGCT 800  
267 uPheAspLeuSerGlyIleLeuThrValPhePheCysGlyIleValMetS 284  
801 TTTGCACTTGACCGGTATCTCCTGCTGCTGCTTTTCTGTTGTTGTTGTT 850  
284 erHisTyrThrTrpHisAsnValThrGluSerSerArgIleThrThriys 300  
851 CCCATTACACATGGCACATCTAAGGAGAGCTCAAGATTAACAAACAAAG 900  
301 HisThrPheAlaThrLeuSerPheLeuAlaGluThrPheIlePheLeuTy 317  
901 CATACCTTTTGAACCTTTGCTATTTCTTCTGCGAGACATTTATTTCTTGA 950  
317 rValGlyMetAspAlaLeuAspIleAspLysTrpArgSerValSerAspT 334  
951 TGTGGAAATGGATGCTTGGACATTTGACAACTGGAGATCCGTGAGTGACA 1000

334 hrProGlyThrSerIleAlaValSerSerIleLeuMetGlyLeuValMet 350  
1001 CACCGGGAACATCGATCGAGTCAGTCAATCCTAATGGTCTGTGTCATG 1050  
351 ValGlyArgAlaAlaPheValPheProLeuSerPheLeuSerAsnLeuAl 367  
1051 GTTGAAGAGCAGCGTTCGTTTCCGTTATGCTTCTATCTAATCTAGC 1100  
367 alyLysAsnGlnSerGluLysIleAsnPheAsnMetGlnValValIleT 384  
1101 CAGAAGAATCAAGCGAGAAATCAACTTTAATGTCAGGTTGTCATTT 1150  
384 rptPsrSerGlyLeuMetArgGlyAlaValSerMetAlaLeuAlaTyrAsn 400  
1151 GGTGGTCTGCTCATGAGAGTGTGTATCTATGCTCTTCATACAAC 1200  
401 LysPheThrArgAlaGlyHisThrAspValArgGlyAsnAlaIleMetI 417  
1201 AAGTTTACAAGGCGCGGCACAGATGTACGCGGAATGCAATCATGAT 1250  
417 eThrSerThrIleThrValCysLeuPheSerThrValValPheGlyMetL 434  
1251 CAGGATGAGATAACTGTCTGCTTTTACACAGTGGTGTGTTGGTATGC 1300  
434 euThrLysProLeuIleSerTyrLeuLeuProHisGlnAsnAlaThrThr 450  
1301 TGACCAACCACTCATAGCTACCTATTACCGCACACAGAACGCCACAG 1350  
451 SerMetLeuSerAspAspAsnThrProLysSerIleHisIleProLeuLe 467  
1351 AGCATGTTATCTGATGACAAACACCCCAAAATCCATACATATCCCTTGT 1400  
467 uaspGlnAspSerPheIleGluProSerGlyAsnHisAsnValProArgp 484  
1401 GGACCAAGACTCGTTCATTGAGCTTCAGGAAACCAATGTGCCTCGGC 1450  
484 roAspSerIleArgGlyPheLeuThrArgProThrArgThrValHistyr 500  
1451 CTGACAGTATACGTGGCTTCTTGACACGCGCCACTCGAACCGTGCATTAC 1500  
501 TyrTrpArgGlnPheAspAspSerPheMetArgProValPheGlyGlyAr 517  
1501 TACTGGAGACAATTTGATGACTCTTCATCGACCGCTCTTTTGGAGTGG 1550  
517 gGlyPheValProPheValProGlySerProThrGluArgAsnProProA 534  
1551 TGCTTTGTACCCCTTTGTTCCAGGTTCTCCAACTGAGAGAAACCCCTCTG 1600  
534 spLeuSerLysAla 538  
1601 ATCTTAGTAAGGCT 1614

seq\_name: gb\_p11:AF056190

seq\_documentation\_block:

LOCUS AF056190 1619 bp mRNA PLN 06-JUN-2000

DEFINITION Arabidopsis thaliana Na<sup>+</sup>/H<sup>+</sup> exchanger (NHX1) mRNA, complete cds.

ACCESSION AF056190

VERSION AF056190.1 GI:5650176

KEYWORDS

SOURCE thale cress.

ORGANISM Arabidopsis thaliana

Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;

Brassicales; Brassicaceae; Arabidopsis.

REFERENCE 1 (bases 1 to 1619)

AUTHORS Quintero,F.J., Blatt,M.R. and Pardo,J.M.

TITLE Functional conservation between yeast and plant endosomal

Na<sup>(+)</sup>/H<sup>(+)</sup> antiporters

JOURNAL FEBS Lett. 471 (2-3), 224-228 (2000)

MEDLINE 20231718

REFERENCE 2 (bases 1 to 1619)



AUTHORS Quintero, F.J.  
TITLE Direct Submission  
JOURNAL Submitted (30-MAR-1998) Biologia Vegetal, IRNASE (CSIC), Avda.  
Reina Mercedes s/n, Sevilla 41012, Spain  
FEATURES Location/Qualifiers  
source 1..1619  
/organism="Arabidopsis thaliana"  
/db\_xref="taxon:3702"  
gene 1..1619  
/gene="NHX1"  
CDS 3..1619  
/gene="NHX1"  
/note="Na+/H+ antiporter"  
/codon\_start=1  
/product="Na+/H+ exchanger"  
/protein\_id="AAF21755.1"  
/db\_xref="GI:6650177"  
translation="M L D S L V S K P L S T S D H A S V A L N L F V A L L C A C I V L G H L L E E N R  
W N E S I T A L L I G L G T V T I L L I S K G S H L L V F S E D L F F Y L L P P I F N A G F O V K K Q  
F R F N F V T I M L F G A V G T I I S C T I I S L G V T O F P K K L D I G T F D L G D Y L A I G A I F A A T D S V C  
T L Q V L N O D E P L L Y S L V F G E V N D A T S V V F N A I Q S F D L T H L N H E A P H L L G N F L Y L  
F L L S T L L G A T G L L S A Y K I R K L F G R H S T D R E V A L M M L M A Y L S T M L A E L D L S G I L T V  
P G T S I A V S S I L M G L V M V G R A A F P P L S F L S N L A K N Q S E K I N F N M Q V Y I W S G L M R G A  
V S M A L A Y N K K T R A C H T D V R G N A I M T S T I T V C L F S T V V F G M L T K P L I S Y L L P H O N A T T  
S M L S D D N T P K S I H T P L L D O D S F I E P S G N H N V P R P D S I R G L F L T R P T R I V H Y Y W R Q F D D S  
F M R P V F G R G E V P P G S P T E R N P P D L S K A "

BASE COUNT 389 a 346 c 360 g 524 t  
ORIGIN

alignment\_scores:  
Quality: 2755.00 Length: 538  
Ratio: 5.121 Gaps: 0  
Percent similarity: 100.000 Percent identity: 100.000

alignment block:

US-09-271-584A-2 x AF056190

Align seg 1/1 to: AF056190 from: 1 to: 1619

1 MetLeuAspSerLeuValSerLysLeuProSerLeuSerThrSerAspHis 17  
|||||  
3 ATCTGGATTCTCTAGTGTGAAACTGCCTTCGTTATCGACATCTGATCA 52  
17 sAlaSerValAlaLeuAsnLeuPheValAlaLeuLeuCysAlaCysI 34  
|||||  
53 CGCTTCTGTGCTGCGTTGAACTCTCTTTGTTGCACATCTTTGTGCTGTA 102  
34 leValLeuGlyHisLeuGluGluAsnArgTrpMetAsnGluSerIle 50  
|||||  
103 TTGTCTTGTGCTATCTTTTGGAGAGAATAGATGGATGAACGAATCCATC 152  
51 ThrAlaLeuLeuIleGlyLeuGlyThrGlyValThrIleLeuLeuIle 67  
|||||  
153 ACCGCTTGTGATTGGGCTAGGCACTGGTGTATACCATTTTGTGATTAG 202  
67 rLysGlyLysSerSerHisLeuLeuValPheSerGluAspLeuPhePheI 84  
|||||  
203 TAAAGGAAAAAGTCGCATCTTCGCTCTTTAGTGAGATCTTTTCTTCA 252  
84 leTyrLeuLeuProProIleIlePheAsnAlaGlyPheGlnValLysLys 100  
|||||  
253 TATATCTTTGCCACCCCATTAATATCAATGCAGGGTTTCAAGTAAAAAG 302  
101 LysGlnPhePheArgAsnPheValThrIleMetLeuPheGlyAlaValG 117  
|||||  
303 AAGCAGTTTTTCCGAATTTTCGTGACTATTATGCTTTTGGTCTGTGG 352  
117 yThrIleIleSerCysThrIleIleSerLeuGlyValThrGlnPhePhe 134  
|||||  
353 GACTATTATTTCTTGCAATATATCTCTAGGTGTAACACAGTCTCTTA 402

134 yLysLeuAspIleGlyThrPheAspLeuGlyAspTyrLeuAlaIleGly 150  
|||||  
403 AGAAGTTGGACATTTGGAAACCTTTGACTTGGGTGATTATCTTCTTATGCT 452  
151 AlaIlePheAlaAlaThrAspSerValCysThrLeuGlnValLeuAsnG 167  
|||||  
453 GCCATATTTGCTGCAACAGATTACAGTGTGTACACTGCAGGGTCTGAATCA 502  
167 nAspGluThrProLeuLeuTyrSerLeuValPheGlyGlyGlyValVal 184  
|||||  
503 AGACGAGACACCTTTGCTTTTACAGTCTTGTATTCGGAGAGGGTGTGTGA 552  
184 snAspAlaThrSerValValPheAsnAlaIleGlnSerPheAspLeu 200  
|||||  
553 ATGATGCAACGTCAGTGTGGTCTTCAACGCGATTACAGAGCTTTGATCTC 602  
201 ThrHisLeuAsnHisGluAlaPheHisLeuLeuGlyAsnPheLeuTy 217  
|||||  
603 ACTCACCTAAACACGAAAGCTGCTTTTCATCTTCTTGGAAACTTCTTGTA 652  
217 rLeuPheLeuLeuSerThrLeuLeuGlyAlaAlaThrGlyLeuIleSer 234  
|||||  
653 TTTGTTTCTCCTAAGTACCTTGTGGTGTGCAACCGGCTCTGATAAGTG 702  
234 laTyrValIleLysLysLeuTyrPheGlyArgHisSerThrAspArgGlu 250  
|||||  
703 CGTATGTTATCAAGAACTATACTTTTGAAGGCACCTCACTGACCGAGAG 752  
251 ValAlaLeuMetMetLeuMetAlaTyrLeuSerTyrMetLeuAlaGluLe 267  
|||||  
753 GTTGCCCTTATGATGCTTATGCGGTATCTTCTTATATGCTTGTGAGCT 802  
267 upheAspLeuSerGlyIleLeuThrValPhePheCysGlyIleValMet 284  
|||||  
803 TTTGACTTGAGCGGTATCCTCAGTGTGTTTTTCTGCTGATTTGTGATGT 852  
284 erHisTyrThrTriphHisAsnValThrGluSerSerArgIleThrLys 300  
|||||  
853 CCCATTACACATGGCACAAATGTACGGGAGAGCTCAAGAAATACACAAG 902  
301 HisThrPheAlaThrLeuSerPheLeuAlaGluThrPheIlePheLeuTy 317  
|||||  
903 CATACCTTTGCAACTTTGTCATTTCTTGGGAGACATTTATTTCTTGTA 952  
317 rValGlyMetAspAlaLeuAspIleAspLysTrpArgSerValSerAsp 334  
|||||  
953 TGTGGAATGGATGCTTTGGACATTTGACAAGTGGAGATCCGTTGATGACA 1002  
334 hrProGlyThrSerIleAlaValSerSerIleLeuMetGlyLeuValMet 350  
|||||  
1003 CACCGGGAAACATCGATCGAGTGAGCTCAATCCTTAATGGGTCTGTCATG 1052  
351 ValGlyArgAlaAlaPheValPheProLeuSerPheLeuSerAsnLeuAl 367  
|||||  
1053 GTTGAAGAGCAGCGCTTCTGCTTTTCCGTTATGCTTCTTCTTAACTTAGC 1102  
367 aLysLysAsnGlnSerGluLysIleAsnPheAsnMetGlnValValIle 384  
|||||  
1103 CAAGAAGAAATCAAGCGAGAAATCAACTTTAACATGCAAGGTGTGATTT 1152  
384 rTrpSerGlyLeuMetArgGlyAlaValSerMetAlaLeuAlaTyrAsn 400  
|||||  
1153 GGTGCTCTGCTCATGAGAGGTGTGTATCTATGCTCTTGCATACAAAC 1202  
401 LysPheThrArgAlaGlyHisThrAspValArgGlyAsnAlaIleMetI 417  
|||||  
1203 AAGTTTACAAGGGCGGGCACACAGATGTACCGGGAATGCAATCATGAT 1252  
417 erThrSerThrIleThrValCysLeuPheSerThrValValPheGlyMet 434  
|||||  
1253 CACGAGTACGATAACTGTCTGCTTTTGTAGCACAGTGGTGTGTGTATGC 1302  
434 euThrLysProLeuIleSerTyrLeuLeuProHisGlnAsnAlaThrThr 450

|||||  
1303 TGACCAACCACTCATAGCTACCTATTACCGCACCAACGCCACACG 1352  
451 SerMetLeuSerAspAsnThrProLysSerIleHisIleProLeuLe 467  
|||||  
1353 AGCATGTTATCTGATGACACACCCCAATCCATACATATCCCTTTGTT 1402  
467 uAspGlnAspSerPheLeGluProSerGlyAsnHisAsnValProArgp 484  
|||||  
1403 GGACCAAGACTCGTTTCATTGAGCTTCAGGCAACCAATGTGCTCGGC 1452  
484 rAspSerIleArgGlyPheLeuThrArgProThrArgThrValHisTyr 500  
|||||  
1453 CTGACAGTATACGGTCTCTTGACACGGCCACTCGAACCGTGCATTAC 1502  
501 TyrTrpArgGlnPheAspSerPheMetArgProValPheGlyGlyAr 517  
|||||  
1503 TACTGGAGACATTTGATGACTCTTCATGCGACCCGCTTTTGAGGTGG 1552  
517 gGlyPheValProPheValProGlySerProThrGluArgAsnProProA 534  
|||||  
1553 TGGCTTTGTACCCTTGTTCTCAGGTTCTCCAACCTGAGAGAACCCCTCTG 1602  
534 spLeuSerLysAla 538  
|||||  
1603 ATCTTAGTAGGCT 1616

seq\_name: gb\_p13:TM021B04

seq\_documentation\_block: 90019 bp DNA PLN 12-JUN-1997  
LOCUS TM021B04  
DEFINITION Arabidopsis thaliana BAC TM021B04.  
ACCESSION AF007271  
VERSION AF007271.1 GI:21911181  
KEYWORDS  
SOURCE thale cress.  
ORGANISM Arabidopsis thaliana  
Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;  
Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;  
Brassicales; Brassicaceae; Arabidopsis.  
1 (bases 1 to 90019)

REFERENCE  
AUTHORS Dante, M., Wamsley, P. and Gibson, A.  
TITLE The sequence of A. thaliana TM021B04  
JOURNAL Unpublished (1997)  
REFERENCE  
AUTHORS Washington University Genome Sequencing Center.  
TITLE The A. thaliana Genome Sequencing Project  
JOURNAL Unpublished (1997)  
REFERENCE  
AUTHORS Waterston, R.  
TITLE Direct Submission  
JOURNAL Submitted (06-JUN-1997)  
COMMENT Submitted by:  
Genome Sequencing Center  
Department of Genetics, Washington University,  
St. Louis, MO 63108, USA  
e-mail: rwlison@watson.wustl.edu

MAPPING: Clones were assigned to the YAC map by hybridization by  
M. Lohdi, Cold Spring Harbor Laboratories, and fingerprinted  
by M. Marra, WashU, to pick the best candidates for sequencing.

NOTICE: This sequence may not be the entire insert of this clone.  
It may be shorter because we only sequence overlapping sections  
once, or longer because we provide a small overlap between  
neighboring submissions.

This sequence was finished as follows unless otherwise noted:  
all regions were double stranded or sequenced with an alternate  
chemistry; an attempt was made to resolve all sequencing problems,  
such as compressions and repeats; all regions were covered by

sequence from more than one subclone

NEIGHBORING COSMID INFORMATION:

The actual start of this clone is at base position 1 of TM021B04 ;  
the actual end is at base position 90019 of TM021B04. The  
orientation of this clone is unknown.

NOTES:

Coding sequences below are predicted from computer analysis, using  
the program Genefinder(P. Green and L. Hillier, ms in preparation).

FEATURES  
source

Location/Qualifiers  
1..90019  
/organism="Arabidopsis thaliana"  
/cultivar="Columbia"  
/db\_xref="taxon:3702"  
/chromosome="V"  
/clone="TM021B04"  
/map="5"  
1802..3685  
/gene="A\_TM021B04.11"  
join(1802..2050,2468..2686,2758..2844,2901..3095,  
3275..3685)  
/gene="A\_TM021B04.11"  
/note="contains similarity to Synecococcus PCC7942  
chromosomal region used as basis of neutral siteir  
recombinational cloning vector (PID:g1174192)"  
/codon\_start=1  
/evidence=not\_experimental  
/protein\_id="AAB61082.1"  
/db\_xref="GI:2191197"  
/translation="MVMLHSVGLIPCSNQKSFLEPHSYRYRCIVCSSETGLSIRRO  
ALEQVDSKSSGDERAALSLVDLQKPDGLRCFGAARQVQRLRYLLEELKLNINFA  
SLSDPTDTTIGSTERNLQAAVSGGIVAWKAEDLSLQQLFTLTGFMFLWLDLVSN  
GGIGSLVLDTTGHTFSQRYHNRVQKHYYIIFHWTYCELSQHEAGHFLVAYLVGLPR  
GYTLSSLEALQKEGSLNIQAGSAFVDYEFLEPNNKLCILLFONQMLNRFSCIALGVA  
TEVLLYCYAPGGDDISKYSFLLPLKNSDDYVNMLYGFVLMEOGLDLVKSIGFTQKK  
ADSQVRWSVLNTILLRRHEIARSKLAQAAMSGESVSGCIQIETSIDPSDI"  
complement(3906..4727)  
/gene="A\_TM021B04.14"  
complement(join(3906..4337,4418..4497,4601..4727))  
/gene="A\_TM021B04.14"  
/note="coded for by A. thaliana cDNA T22079"  
/codon\_start=1  
/protein\_id="AAB61076.1"  
/db\_xref="GI:2191191"  
/translation="MEATSLSSAATISSSSPLSIFSPKRTDSPPPRIVRLSNKK  
EDKDYDPQHSNSSLFRNRTLSNDEAMGLVLSAASVKGWTTSGMGPSLPKATDT  
DIVSTFWSLFTKSPRRMRVAFTCNVCGORTTAINHAYTDGTVFVQCCGCVFHK  
LVDNLNLFHEVAYVSSSFDYDAKWVDSGLNLFDEDDDDNAGDSNDVFPPL"  
complement(12650..16625)  
/gene="A\_TM021B04.10"  
complement(join(12650..13129,13415..13509,14187..14498,  
14596..16014,16127..16625))  
/gene="A\_TM021B04.10"  
/note="similar to N. tabacum membrane-associated  
salt-inducible protein (PID:g473874)"  
/codon\_start=1  
/evidence=not\_experimental  
/protein\_id="AAB61077.1"  
/db\_xref="GI:2191192"  
/translation="MKSDFLTSTTHNPISFIPKIPSRNSRLISKSSSSSKVRPDPP  
SLSDGNPEKPKPRYEPKPHISDDARRLIKKKAYLSLRLRNQSQAMTPKWIKRTTP  
EQWQVLEDDRQMYGKRVAAIKTVRGLSRQSGDDMRVMSFVAKUSFDMQVACG  
VLKEQRCFCVKQLSVPVSVVYTVILRYLGQVGIKMAEETFLEMVECEPDVACG  
TMCITYARWRHSAMLTFFYKQERRILLSTSVNFMSSLQKSFHGKVIDLWLEW  
EEGVPNFEYTLVSSYAKQGFKEALKAFGEMKSLGFVPEVTVSSVISLSVAGD  
WKAIGLYEDMRSGQVPSNYTCATMLSYKYTENYKALSLFADMERKNIPADDEVIR  
GKILRYGLGLFHDQASMEETERNLLADEKTYLAMSOVHLNSGNDMLNVDIEMM  
KTRDIPLSRFATVIMLQCYAKIQNVDAEEAFRLSKTGLPDPASSGNDMLNVDIEMM

gene  
CDS  
complement(17345..24723)  
/gene="A.TM021B04.9"  
complement(join(17345..17911,17988..18340,19797..20049,20117..20610,21561..24723))  
/gene="A.TM021B04.9"  
/note="contains similarity to a DNAJ-like domain"  
/codon\_start=1  
/evidence=not\_experimental  
/protein\_id="AAB61072.1"  
/db\_xref="GI:2191187"  
/translation="MDWNKEACRAKTLAEDKMKEGDFVGAQKLLKLAQSLFSGLES  
PQMLAVCDVHNSAEKKINCLENWYGIQVHMFADDAVIKKQVRKLAALLHPDKNOFPG  
AEAPFLVMDASRLADKDRSOYDIRRIYLRLATNOLNANSGLQCAATNSADTFW  
TCCEHCGYRYKLYBKLYNILLNCILQORSYMYDTGNEAPSKSNTQKQEVONGPCN  
TSLNTNGSIGAQPQSAEAVDDKGTNKNKNGGSGESKTEVKNSKIEETEVRND  
DEEKMSAKLHKHPHPEVEPEIGASKSPVDESRSDEAPSTSKDNKKKGVYESI  
NVGDSDFIKDAYSKDNKKRSPRSQSSYAEKISDNLGPPKKRLKSLNVGLKSE  
QTRKGLVGVGGSKRFDSSGSAFCAFKRKAFFVDSGYOEISSVKDEREVRKAS  
GEGVMEAKIDNNHNPENLITDLPPEFSNFEITWSCFGVNOVMSYDPIDGMPRL  
YARIDKVLVPEFKLWIDPLQDNKNSIPFIAGITQGGGSEENHDLFSCOMFHL  
TRNNSVIYPRKGEIWAIFRGWDISWASSENHHPHPEYDEVELSNFNDENGVLGVF  
LGKGVFSVLPQDAQGVQLQIPSSOMLRFHSKVPFKMTGKEREGVPTGCFELPD  
AALPKFELVDKVDVGLDELPGKGTGPPPEASKVEMOANTRESSPPKKRPKV  
HNDHGSFSGSDGRTHERNASKEKRSKVADALNLKSPRLSLSNSQANSRSGR  
QIAEKKSANNGSCGKPGGLSISGKMSKSLITRMKTPQDIHKPTVNVKKHGRNDS  
LSQSGDGLQLNGNMKSSPETRVPSSCKTENTFENQNSWDRKFDQIDWAIY  
SNDSGRPKYAIKIDITSEFKLRVAPLELYRPIIMRPVCCGRFKLTKGAEVTV  
PSSSHQVAKVYKTRFEVYPGKEIWALYKNCNTRDYETELEIVEVYVETQRI  
QAMTLTAQFNKPKLYRRSESNASFIMERNLCEDDSDVDIMLLVVGVSVDAYESOI  
PRPMRQITLTKGHAYIOKALKDDPITFEROLYRMPNPEFAELCHLLOKMTGLKGPVHC  
VEEMVATFLITVGONSRYCHTMDTFKSKFSTSINFKVLPAALNLAFTLMKAVNTVY  
PSKTSKTRFEYYPKDCVGAIDGTHINMWOGPEKASYRNKRGVISONVLAACNFLE  
FIYVLSWEGSAHDSKVLDQALTRTNKRLQVPEGNRLQVLKGTITMGDSQKRRKKGDY  
NPWSPETKLLVLLVEGINNNRDSNGTISKLVETKFMPEINKEFCRCKNNHYLS  
RMKLYIQOQSCDLQRFSGGWDPLTKRFTASDEVMSDYLKAPNNKQLRYDTFEF  
FDELIQIFEGVATGKNAIGLSDCTDGLTYRAGENPRKEYVDDFDNVYEDVTTTHES  
SEHVAPMNSHTSSESKPLPRKTRSERSTSOKEESPMMVSSKILLDIIQOREBQON  
EVAQKNMWDAIEISDLDECIRYKALTKIYHLGIDQOVFVSNVSVERLQWQITIME"  
complement(34437..37759)  
/gene="A.TM021B04.8"  
complement(join(34437..35003,35094..35272,35446..36036,36081..36197,36313..36607,36724..36905,37102..37759))  
/gene="A.TM021B04.8"  
/note="contain similarity to type 1 inositol  
1,4,5-trisphosphate receptors"  
/codon\_start=1  
/evidence=not\_experimental  
/protein\_id="AAB61078.1"  
/db\_xref="GI:2191193"  
/translation="MKAKQSEFERKEKDFLEQKAEVEKKRREVEQLEKFTTRMESVE  
RVSDEKLMELGLRATLELELQKMEVEKRRERIVAGDKLRGEFPLVLSLAKNMGLSVTM  
PVKSTVLENNADVMKNTALARMVYLPDPAKVLDIAEGSFKEYWKXDLGEADDR  
VWNSVILLENLILKNLKITPQVQOEATPLGIAWLGRKANMKNDPPQVGCALFLAA  
YGAVENTLKKIEYLATKFIICEFRLYKLCPPGRGPELLIEFFDSSDKAARVIAGTGTS  
MEAQKARREKKKADAMAIIKIEAKAETMPPAKILKRLAVRNDESAQRAMEPQVKS  
TEKQSTTKGVEKSEAKSPIYEQKHVIKRPRLTEPTAPNLTVCFNKMHKELTSYM  
NLSIFVGETLRHFNQNIIFLLIKQPEVVCPTQVKQKESGADHQPTIATHPSG  
TETNLISGSIKADMLRELVEKPLKESDLSNALKCTPDPAKLFIDLTSMALCPTWT  
EGGYEFLKLTISASCLLNOLKLLPKIGHPVKGDAKKLAVYWKDKIAKSRDOLVY  
ICFLQFLGIFGIVSEFKADLLGLDLSNYQTVSPDLCQLGLDDAIPGFQNLKRGV  
HRIKADYIYFSGVHRPQPSVAILNSLRTIKESAESKSYREAKNSTDVQVAILDRGV  
RALRAALKEISCHLESFQLGDLDEQIKSLLLKRLRNTSNGSGSASPDSTIKOS  
QAKPPTVAYEAPVYTSNIPLEPTEAASASSKPFSSKNRGRKRSNGNSQSGHIA  
SHTSNHYPSHDYSLNQLRTPVDNDRGFTGFPPNDYNNQWQPEGPPQFVHYLQPLD  
PRYRN"

gene  
CDS  
complement(39523..43518)  
/gene="A.TM021B04.7"  
complement(join(39523..39921,40013..40173,40283..40858,41109..43518))  
/gene="A.TM021B04.7"  
/note="contains similarity to tropomyosin and kinesin"  
/codon\_start=1  
/evidence=not\_experimental  
/protein\_id="AAB61071.1"  
/db\_xref="GI:2191186"  
/translation="MBEIKLENEIRLCDVKANNIRKTMDIKMSQASDVLIILNQWCFD  
alignment\_scores:  
Quality: 2282.50 Length: 1007  
Ratio: 4.250 Caps: 18  
Percent Similarity: 53.327 Percent Identity: 53.029  
alignment\_block:  
US-09-271-584A-2 x TM021B04/rev ..  
Align seg 1/1 to reverse of: TM021B04 from: 1 to: 90019  
1 MetLeuAspSerLeuValSerLysLeuProSerLeuSerThrSerAspHi 17  
68024 ATGTTGGATTCTCTAGTGTGCAAACTGCCTTCGTTATCGACATCTGATCA 67975  
17 salaserValValAlaLeuAsnLeuPheValAlaLeuLeuCysAlaCysI 34  
67974 CCCTCTCTGTGGTGGTGAATCTCTTTGTTGACATCTCTTTGCTGTTGTA 67925  
34 leValLeuGlyHisLeuLeuGluAsnAtgTTrpMetAsnGluSerIle 50  
67924 TTGTTCTTGGTCATCTTTTGGGAAGAATAGATGATGAACGAATCCATC 67875  
51 ThrAlaLeuLeuIle..... 55  
67874 ACCGCTTGTGATTGTTAGTTCACTCAATGCAAAAGAGTAATTACTGAG 67825  
56 .....GlyLeuGlyThrG 60  
67824 TTCCAACCTTGAATGACATTTGAGCTATTGTTGTTGCAAGGCGGTAGCAGCTG 67775  
60 lyValThrIleLeuLeuLeuSerLysGlyLysSerSerHisLeuVal 76  
67774 GGTGTACCACTTTGTTGATTAGTAAGAAAGTCGCATCTCTCTCTGTC 67725  
77 PheSerGluAspLeuPhePheIleTyrLeuLeuProIleIlePheAs 93  
67724 TTTAGTGAAGATCTTTCTTCATATATCTTTTGCCACCCATATATTCAA 67675  
93 nAla.Gly..... 95  
67674 TGCAGGGTCAGTCTTTTTCATCATCCATTTGAGCCCTTTTCCTCTATGT 67625  
95 ..... 95  
67624 TGGGATATATCTTTTAATGTAGGACTTAGTTGGTAGGCTCTCATGCTAT 67575  
95 ..... 95  
67574 ATGATACGTTTAGTTGCACGGGTTTACCTACAAAAGTTACATAAAACAG 67525  
95 ..... 95  
67524 ATTGAAGAAAAAACTTAGTTTGAAGAATGATGCATTTATGTGATTCTAGG 67475  
95 ..... 95  
67474 TGATAAAGTTAGTACAACGCTTTTAAAGATGTTGATTAGTAAGCTTAAAGCT 67425  
95 ..... 95  
67424 CTACTTCTTAATAATACCTGTGATACATCAACTTTTCTTAGTACTTGTGAT 67375

```

95 ..... 95
67374 TGTATAAAGAAAGAGGCATGATGCTCTTTTGAAGTAGTAGATT 67325
95 ..... 95
67324 ATACACGTGCAAGTAAGTGTGGTCTGGATTTCTCTTTCATGGATCTT 67275
96 ..... PheGlnValLysLysGlnP 103
67274 TCTCAGATTGTTGTGTCTATTTTCAGGTTTCAAGTAAAGAACAGT 67225
103 hePheArgAsnPheValThrIleMetLeuPheGlyAlaValGlyThrIle 119
67224 TTTTCCGCAATTTTCGACATATATATGCTTTTGGTCTGTGGACATTT 67175
120 IleSerCysThrIleLeuSerLeu..... 127
67174 ATTTCTTGCACAATCATATCTTAGGTAGCTTACTTCGCTTTACGATTGG 67125
127 ..... 127
67124 CTGAACACCCGCTGAATGCGGCTTTTGGATTAGTCGAGATAGTTTGT 67075
128 ..... GlyValThrGlnPhe 132
67074 ATATCTCTACTGATTTTCTGGTGGTCTTGAACACAGGTGTAAACACAGTTC 67025
133 PheLysLysLeuAspIleGlyThrPheAspLeuGlyAspTyrLeu..... 147
67024 TTTAAGAAGTTGGACATTTGAACCTTTGACTTGGTGATTATCTTGTTAA 66975
147 ..... 147
66974 GGCCTGAACTAATGGCTTTCTTTTGGTTCCTATTTGCTTTACTTAATCGGG 66925
148 ..... AlaIleGlyAlaIlePhe 153
66924 AGACTTGATTTGTTTGGTGTGTGANTCTCAGCTATTTGGTGCATATTT 66875
154 AlaAlaThrAspSerValCysThrLeuGlnVal..Leu..... 165
66874 GCTGCAACAGATTCAGTATGTACACTGCAGGTAATGTGTTTCTCCAACA 66825
165 ..... 165
66824 TGACTTAAATTACTTACAAACCCCTGAACACATAGTGGACCGAAAGTTTCTG 66775
166 ..... AsnGlnAspGluThrProLeuL 173
66774 TTCTGTGTGTTGTGATATAGGTTCTCGAATCAGAGCAGACACCTTTGC 66725
173 euTyrSerLeuValPheGlyGluGlyValValAsnAspAlaThrSerVal 189
66724 TTTACAGTCTGTATTCGGAGAGGCTTCTGTAATGATGCAACGTCAGTT 66675
190 ValValPheAsnAlaIleGlnSerPheAspLeuThrHisLeuAsnHisG1 206
66674 GTGGCTTCAACGCGATTTCAGAGCTTTGATCTCCTCACCCTAAACCCACGA 66625
206 uAlaAlaPheHisLeuLeuGlyAsnPheLeuTyrLeuPheLeuLeuSerT 223
66624 AGCTGCTTTTCATCTTCTGGAACACTTCTGTATTTCTTCTCCTAAGTA 66575
223 hrLeuLeuGlyAla..Ala..... 228
66574 CTTTGGTGTGGTGCAGTAAGTCCATTGTTTCTTCTCCCTGCCATTTCTTT 66525
228 ..... 228
66524 CTTGTAGCTGCTATATGACGCTCTGCTGTGTTCATATCTTTGTTATTCATC 66475

```

```

229 .....ThrGlyLeuIleSerAlaTyrValIleLysLysLeuTyrPheG1 243
66474 TTGCAGACCGGCTCTGATAAGTGCATGTTTATCAAGAAAGCTATATACTTTGG 66425
243 y..... 243
66424 AAGGTACATATTTTAACGTCAATATATGTGACTTGTCTTTTAAAGTTGTTT 66375
244 ..... ArgHisSer 246
66374 CCGTGGTCTAAGTAGAGTGGCTCACATTAATGTACAGGCACATCA 66325
247 ThrAspArgGluValAlaLeuMetMetLeuMetAlaTyrLeuSerTyrMe 263
66324 ACTGACCCGAGAGGTTGCCCTTATGATGCTTATGGCGTATCTTTCTTATAT 66275
263 tleuAlaGluLeu..... 267
66274 GCTTCTGAGGTTAGATTCTTTTAAATAACACGAACATATTTTCAGTT 66225
268 ..... Phe 268
66224 GCATTCAACACAAATATATATGTATATTAATTTGGTTGCAGCTTT 66175
269 ..AspLeuSerGlyIleLeuThrValPhePheCysGlyIleValMetSer 284
66174 TCGACTTGAGCGGTATCTCCTCACTGTGTTTTTCTGTGGTATTGTGATGCC 66125
285 HisTyrThrTrpHisAsnValThrGluSerSerArgIleThrThr..... 299
66124 CATTACATGCGCACAAATGTAAACGAGAGCTCAAGATAACAAACAAAGTA 66075
299 ..... 299
66074 TCCAAACTTTAGTTTCTCGGCTTTTGTGAGATTAACACTTTTACCTC 66025
300 ..... Lys..HisThrPheAlaT 305
66024 CCAATGATTATCTTATGATTTTTTCGATGAACAGGCATACCTTTGCAA 65975
305 hrLeuSerPheLeuAlaGluThrPheIlePheLeuTyrValClyMetAsp 321
65974 CTTTGTCAATTTCTCGGAGACATTTATTTCTGTATGTTGGAATGGAT 65925
322 AlaLeuAspIleAspLysTrpArgSerValSerAspThrProGlyThrSe 338
65924 GCCTTGGACATTGACAAGTGGAGATCCGCTGAGTGACACACCGGGAAACATC 65875
338 rIleAlaValSerSerIleLeuMetGlyLeuValMetValGlyArgAlaA 355
65874 GATCGCAGTGAGCTCAATCCTAATGGTCTGGTCATGGTTGGAAGAGCAG 65825
355 laPheValPheProLeuSerPheLeuSerAsnLeuAlaLysLysAsnGln 371
65824 CGTTGCTCTTTCCGTTATCGTTTCTATTAAGTCTAGCCCAAGAAAGTCAA 65775
372 SerGluLysIleAsnPheAsnMetGlnVal..Val..... 382
65774 ASCGAGAAATCACTTTAACATGAGGTATGATCATCTACTTCTGTTAC 65725
382 ..... 382
65724 TAATAGATTCAATAATTTCTCTGGATTTTCAGCTATTTATATCTATCAATTT 65675
383 ..... IleTrpTrpSerGlyLeuMetArgGlyVal 392
65674 TTTCCTCATTTTCTTAGGTTGTGATTTGGTGGTCTGCTCATGAGAGGTGC 65625
392 aValSerMetAlaLeuAlaTyrAsnLys..... 401
65624 TGTATCTATGGCTCTTGCATACAAACAGGTATAAAAGAAATGCTCTTTAA 65575
401 ..... 401

```

```

65574 TACTCTCTATGAAGCTTCAAAATGACAAATCATCATTTAAACACTGTGTT 65525
402 .....PheThrArgAlaGlyHisThrAspValArgGlyAsnAlaI 415
65524 CTATTGTAGTTTACAAGGCGGGCACACAGATGTACGCGGAATGCAAT 65475
415 eMetIleThrSerThrIleThrValCysLeuPheSerThrVal..... 429
65474 CATGATCAGAGTACGATCACTGTCTGTCTCTTTTATGACAGTGGTGAAGA 65425
429 ..... 429
65424 ATCTCTGGATGATCAAAATGAGTTACATAATATTTCTTAATGAATGGAC 65375
430 .....ValPheGlyMetLe 434
65374 ATTTGGTTAATAGATCTCTTGAACCTTGTTCAGGTTGTTGGTATGCT 65325
434 uThrLysProLeuIleSerThrLeuLeuProHisGlnAsnAlaThrThrS 451
65324 GACCAACACCATCATAAAGTACCTATTACCGCACCAAGCCACCAACGA 65275
451 eMetLeuSerAspAspAsnThrProLysSerIleHisIleProLeuLeu 467
65274 GCATGTTATCTGATGACACACCCCAAAATCCATACATATCCCTTTGTTG 65225
468 AspGlnAspSerPheIleGluProSerGlyAsnHisAsnValProArgPr 484
65224 GACCAAGACTCGTTTCATTGAGCTTTCAGGGAACCAATGTGCCCTCGGCC 65175
484 oAspSerIleArgGlyPheIleuThrArgProThrArgThrValHisTyrT 501
65174 TGACAGTATACGTTGCTCTTGACAGGCGCCACTCGAACCGTGCATTA 65125
501 yTrpArgGlnPheAspAspSerPheMetArgProValPheGlyGlyArg 517
65124 ACTGGAGACAATTTGATGACTCTTCATCGACCGCCCTCTTTGGAGTCTG 65075
518 GlyPheValProPheValProGlySerProThrGluArgAsnProProAs 534
65074 GGCTTTGTACCTTTGTTCCAGGTTCTCCAACTGAGAGAAACCCCTCTGA 65025
534 pLeuSerLysAla 538
65024 TCTTAGTAAGGCT 65012
seq_name: gb_p11:AB038492
seq_documentation_block:
LOCUS AB038492 2637 bp mRNA PLN 18-AUG-2000
DEFINITION Atriplex gmelini AgNHX1 mRNA for Na/H antiporter Nhx1, complete cds.
ACCESSION AB038492
VERSION AB038492.1 GI:9857313
KEYWORDS Na/H antiporter Nhx1.
SOURCE Atriplex gmelini cDNA to mRNA.
ORGANISM Atriplex gmelini
Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Caryophyllales; Caryophyllaceae; Chenopodiaceae; Atriplex.
REFERENCE
AUTHORS Hamada,A., Hayakawa,T., Shono,M., Xia,T., Hayashi,Y. and Tanaka,A.
TITLE Isolation and characterization of a Na/H antiporter gene from the halophyte Atriplex gmelini
JOURNAL Unpublished (2000)
REFERENCE
AUTHORS Hamada,A., Hayakawa,T., Hamada,A. and Tanaka,A.
TITLE Direct Submission
JOURNAL Submitted (17-FEB-2000) to the DDBJ/EMBL/GenBank databases.
AUTHORS Takahiko Hayakawa, Plantech Research Institute, Research Center;
TITLE 1000 Kamoshida-cho, Aoba-ku, Yokohama, Kanagawa 227-0033, Japan
JOURNAL (E-mail:pr10012@cc.m-kagaku.co.jp, Tel:+81-45-963-3520,

```

```

FEATURES
    source          1..2637
                    /organism="Atriplex gmelini"
                    /db_xref="taxon:118074"
    gene            682..2349
                    /gene="AgNHX1"
    CDS             682..2349
                    /gene="AgNHX1"
                    /codon_start=1
                    /product="Na/H antiporter Nhx1"
                    /protein_id="BAB11940.1"
                    /db_xref="GI:9857314"
                    /translation="MMSQLSSLLSGKMDALTTSDHASVSVSMNLFVALLGCGCIVIGHLL
                    ENRMNESITALLIGLATGVWILLISGGKSHLLVFSDELDFEYILLPPIIFNAGFQV
                    KKQFFNFITIVLFAGVGLVFTIISLGALSIFKKLDITGLLEADYLAIGAIFAAT
                    DSVCTLOVNODETPLLYSLVFGVGVNDATSVLFNATQSPDLTRIDHRIALQPMGN
                    FLXLFIASTILGAFGLLSAVIILKLYFGRHSDREVALMMLMAYLSYMLAEFLYLSG
                    ILTVFPGCIYMSHYTHWNTESRVTTHKAFATLSFAVSEVLEFLVYVMDALDEKWRP
                    VDSPGISVAVSSILGLVWVGRAAFVFPUSLNMENFAKKSQSEKVTFFNOQIVLWNRGL
                    MRGASMALAYNOFTRSGHTOLRGNAIMITSTISVLFSTWVEGLLTKPLIMFLDPOP
                    KHFTSCSTVSDVSGPKSYSLPILLEGNODYEVDVGNHEDTTEPTIVRPSSLRMLLN
                    APHTVHHYWRKFDSEMRPVFGGFGFVPGSPTEQSTNNLVDR"
    polyA_site      2607
                    /note="30 a nucleotides"
BASE COUNT        685 a 509 c 530 g 913 t
ORIGIN
alignment_scores:
    Quality: 2168.50      Length: 547
    Ratio: 4.286          Gaps: 5
    Percent Similarity: 92.505    Percent Identity: 77.697
alignment_block:
US-09-271-584A-2 x AB038492 ..
Align seg 1/1 to: AB038492 from: 1 to: 2637
2 LysAspSerLeuValSer...LysLeuProSerLeuSerThrSerAspHis 17
694 TTAAGCTCTTTTGTGTCTGGAAAGATGGACGCTCTTACCACCTTCTGATCA 743
17 sAlaSerValValAlaLeuAsnLeuPheValAlaLeuLeuCysAlaCysI 34
744 CGCTCTGTGCTCGATGAACCTGTTTCTGGCACGTGTTATGCTGTGTA 793
34 leValLeuGlyHisLeuLeuGluAsnArgTrpMetAsnGluSerIle 50
794 TCGTAATTGGTCTATCTTAGAGGAGAAATCGTTGGATGAATGAGTCCATC 843
51 ThrAlaLeuLeuIleGlyLeuGlyThrGlyValThrIleLeuLeuIleSe 67
844 ACTGCCCTTCTTATAGTTTGGCTACTGGGTGTGATTCGCTGCTGATTAG 893
67 rLysGlyLysSerSerHisLeuValPheSerGluAspLeuPhePheI 84
894 TGGAGGAAAAAGTTCACATCTTTGGTCTTCAGTGAAGATCTTTCTTCA 943
84 leThrLeuLeuProProIleIlePheAsnAlaGlyPheGlnValLysLys 100
944 TATACCTTCTTCCACCATATATTATTCAGGAGGCTTTCAGGTGAGAGAG 993
101 LysGlnPhePheArgAsnPheValThrIleMetLeuPheGlyAlaValG 117
994 AAGCAGTCTTCCCAACTTCATTACTATTGTTGAGCTGTTGG 1043
117 yThrIleIleSerCysThrIleIleSerLeuGlyValThrGlnPhePhe 134
1044 TACATGGTATCATTCACCATCATATCTCTGGGAGCGGTGTGCAATTTT 1093
134 yLysLeuAspIleGlyThrPheAspLeuGlyAspTyrLeuAlaIleGly 150

```

```
1094 AGAAATTGGATATTGGTACTCTGAGCTGGCAGACTATCTTGCAATTGGT 1143
151 AlalPheAlaIaThrAspSerValCysThrLeuGlnValLeuAsnGlu 167
|||||
1144 GCAATATTGCTGCGACAGATTCTGTTGTCACACTGCAGGTTCTTAATCA 1193
167 nAspGluThrProLeuLeuTyrSerLeuValPheGlyGluGlyValVala 184
|||||
1194 GGATGAGACCCCTCTGCTACAGCTGGTCTTTGGCGAGGGTGTGTTA 1243
184 snAspAlaThrSerValValValPheAsnAlaIleGlnSerPheAspLeu 200
|||||
1244 ATGATGCCACATCAGTGGTCTTTTCATGCATTTAGACCTTTGACCTC 1293
201 ThrHisLeuAsnHisGluAlaAlaPheHisLeuLeuGlyAsnPheLeuTy 217
|||||
1294 ACAAGAATTGATCACAAGTAGCTTTACAATTTATGGCAACTTCTTATA 1343
217 rLeuPheLeuLeuSerThrLeuLeuGlyAlaAlaThrGlyLeuIleSerA 234
|||||
1344 TTTATTATCGCAAGCAGCAGTACTTGGAGCATTTACTGGCTTGCTCAGT 1393
234 laTyrValIleLysLeuTyrPheGlyArgHisSerThrAspArgGlu 250
|||||
1394 CTTACATTATCAAAAGCTGTACTTTGGAAGGCATTCCTGATGCTGGAG 1443
251 ValAlaLeuMetMetLeuAlaTyrLeuSerTyrMetLeuAlaGluLe 267
|||||
1444 GTTGCTTTAATGATGCTATGGCTTATCTATCTTACATGCTGCTGAAC 1493
267 upAspLeuSerGlyIleLeuThrValPhePheCysGlyIleValMets 284
|||||
1494 TTTCTATTGAGTGAATCTTACTGCTGATTTCTCTGGGATTCATGT 1543
284 erHisTyrThrTrpHisAsnValThrGluSerSerArgIleThrLys 300
|||||
1544 CCCATTATACCTGGCACAATGTGACAGAGACTCAAGAGTAACCAACA 1593
301 HisThrPheAlaThrLeuSerPheLeuAlaGluThrPheIlePheLeuTy 317
|||||
1594 CATGCTTTTGAACACTATCTTTTGTGCTGAGGTTTCTCTATTTCATA 1643
317 rValGlyMetAspAlaLeuAspIleAspLysTrpArgSerValSerAsp 334
|||||
1644 TGTGCTATGGATGATGGACATTGGAAGTGGAGATTGTTAGTGATA 1693
334 hrProGlyThrSerIleAlaValSerSerIleLeuMetGlyLeuValMet 350
|||||
1694 GTCCTGGCATTTCTGCTGTGAGTTCATATGCTAGGCTAGTCATG 1743
351 ValGlyArgAlaAlaPheValPheProLeuSerPheLeuSerAsnLeuAl 367
|||||
1744 GTTGAAGAGCAGCTTTGTTTTCCTCTATCTGTTGATGAACCTTGC 1793
367 alyLysAsnGlnSerGluLysIleAsnPheAsnMetGlnValValIleT 384
|||||
1794 CAAGAAATCGCAAGTGAAGAGTCACTTTCAACGACGAGATTGTCATAT 1843
384 rpTrpSerGlyLeuMetArgGlyAlaValSerMetAlaLeuAlaTyrAsn 400
|||||
1844 GTGGGCTGGTCTTATGAGAGTGTGTTTCCATGCGACTTGTCTATAAT 1893
401 LysPheThrArgAlaGlyHisThrAspValArgGlyAsnAlaIleMetIl 417
|||||
1894 CAGTTTACGAGCTTGGGCACACAGCTTAGGGGAAATGCAATCATGAT 1943
417 eThrSerThrIleThrValCysLeuPheSerThrValValPheGlyMetL 434
|||||
1944 CACAAGCATAATATCTGTTGCTTTTAGTACAATGGTGTGGGGTGC 1993
434 euThrLysProLeuIleSerTyrLeuLeuProHisGlnAsnAlaThrThr 450
|||||
1994 TGCAAGAGCTCTCATCATGCTTTTGTGCTTCCTCGCTCAACCGAACACTTCACT 2043
```

```
451 .....SerMetLeuSerAspAspAsnThrProLysSerIleHisIlePr 465
|||||
2044 AGTTCACACACCGATCATGATGTGGAAGTCCAAAGTCATACTCGTTGCC 2093
465 oleuLeuAsp.....GlnAspSerPheIleGlu....ProSerGlyAsnH 479
|||||
2094 ACTCCTTGAGGGCAACCAAGATTATGAAGTTGATGTGGGAAACGAAACC 2143
479 is.....AsnValProArgProAspSerIleArg 488
|||||
2144 ATGAAGACACCACTGAGCGCGGACTAGTTTCGACCTAGTAGCCTCGC 2193
489 GlyPheLeuThrArgProThrArgThrValHisTyrTyrTrpArgGlnPh 505
|||||
2194 ATGCTTCTAATATGCACCTACTCACACCGTCCATCACTATTGGCGCAANT 2243
505 eAspAspSerPheMetArgProValPheGlyGlyArgGlyPheValProp 522
|||||
2244 CGATGACTCTTTCATGCGCGCGTGTGTTGGTGGCGGGTGTGTACCTT 2293
522 heValProGlySerProThrGluArgAsnProProAspLeu 535
|||||
2294 TTGTCCTCGGGTTTCACCTACTGAACAAGCACCACCAATTTG 2334
seq_name: gb_pl1:AB033989
seq_documentation_block: 2218 bp mRNA PLN 06-OCT-2000
LOCUS AB033989
DEFINITION Ipomoea nil Purple mRNA for Na+/H+ exchanger, complete cds.
ACCESSION AB033989
VERSION AB033989.1 GI:10716128
KEYWORDS
SOURCE
ipomoea nil (strain:KKZSK-2) fully colored flower buds cdna to
mRNA, clone_lib:cdna library.
ORGANISM
Ipomoea nil
Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
Magnoliophyta; eudicotyledons; core eudicots; Asteridae; euasterids
I; Solanales; Convolvulaceae; Ipomoea.
REFERENCE
1 (sites)
AUTHORS Fukuda-Tanaka,S., Inagaki,Y., Yamaguchi,T., Saito,N. and Iida,S.
TITLE Colour-enhancing protein in blue petals
JOURNAL Nature 407, 581 (2000)
REFERENCE 2 (bases 1 to 2218)
AUTHORS Fukuda-Tanaka,S., Inagaki,Y., Yamaguchi,T. and Iida,S.
TITLE Direct Submission
JOURNAL Submitted (26-OCT-1999) to the DDBJ/EMBL/GenBank databases.
Yoshishige Inagaki, National Institute for Basic Biology, Division
of Gene Expression and Regulation 1; Myodaiji-cho, Okazaki, Aichi
444-8585, Japan (E-mail:yinagaki@nibb.ac.jp, Tel:81-564-55-7684,
Fax:81-564-55-7685)
FEATURES
Location/Qualifiers
source
1..2218
/organism="Ipomoea nil"
/strain="KKZSK-2"
/db_xref="taxon:35883"
/clone_lib="cdna library"
/tissue_type="fully colored flower buds"
300..1928
/gene="Purple"
300..1928
/gene="Purple"
/codon_start=1
/product="Na+/H+ exchanger"
/protein_id="BAB16380.1"
/db_xref="GI:10716129"
```

gene

CDS

/product="Na+/H+ exchanger"

/protein\_id="BAB16380.1"

/db\_xref="GI:10716129"

/translation="MAFGSLQLQNSDLFTSDHASVSMNLFVALLCACIVLGHLEE

NRWNESTITALIGCTGWTLLSGKSHLLVPSDFLYLLPPIFNAGFOVKR

KOFFVNEWTIMLFGAIGTLISCSISFGAVKIFKHLOIDFLDFDGLAIGAIFATDS

VCTLOVLSODETPLLYSLVFGGVNDATSVLENAIQSFDMTSFDPKIGLHFTGNFL

YLFSLSTFLGVGIGILLCAIILKKLYFGRHSTDREVALMLMSLYSLIMAEFLYSGL

TVFFCGIVNHSYTHWNVTSSRKVTRFATLSFVAETFIPLYVGMADALDEKWKFKV

NSQGLSVAVSSILVGLILVGRAAFVFPFLSLNSLAKNSSDKISFRQOIIILWAGLMR

GAVALAYNKFTTSGHTSLHENAIMITSTVTVVFLSTVVFGLMTRKPLINLLPPPHKQ  
MPSGHSNMTSESPSKHFTVPLIDNPDSESDMITGPEVARPTALRMILLRTPTHTVH  
RYWRKFDSEMRPVPFGGRGFVPFVAGSPVQSPR"  
BASE COUNT 568 a 439 c 492 g 719 t  
ORIGIN

alignment\_scores:  
Quality: 2041.00 Length: 541  
Ratio: 4.182 Gaps: 4  
Percent Similarity: 90.203 Percent Identity: 73.752  
alignment\_block:  
US-09-271-584A-2 x AB033989 ..  
Align seg 1/1 to: AB033989 from: 1 to: 2218  
2 LeuAspSerLeuValSerLysLeuProSerLeuSerThrSerAspHisAl 18  
312 TTGCTCTCTTCTGCTCAAAAT...TCGGATTGTTACGCTCGATCATGC 358  
18 aSerValValAlaLeuAsnLeuPheValAlaLeuLeuCysAlaCysIleVal 35  
359 TTCGGTTGTCGATGAACCTCTTTGTCGCGFTGCTTTCGCGCATGCATTG 408  
35 aLeuGlyHisLeuLeuGluGluAsnArgTTPMetAsnGluSerIleThr 51  
409 TTTCTGGCCCATCTACTCGAGGAGATCGCTGGGTGAACGAATCCATTACT 458  
52 AlaLeuLeuIleGlyLeuGlyThrGlyValThrIleLeuLeuIleSerLys 68  
459 GCCCTTATAATTTGGTTGTGTCACCGGAGTTGTAATTTTGCCTTTCAGCGG 508  
68 sGlyLysSerSerHisLeuLeuValPheSerGluAspLeuPheIleThr 85  
509 AGGAAGAGTTCCATCTCTCTGCTTTAGCGAAGATCTTTTCTTTATAT 558  
85 yrLeuLeuProIleIlePheAsnAlaGlyPheGlnValLysLysLys 101  
559 ATCTCTGCGCCATATATATATTCATGCGGGTTTCAAGTGAAGAAGAG 608  
102 GlnPhePheArgAsnPheValThrIleMetLeuPheGlyAlaValIleGly 118  
609 CAGTTTTCGTGAACCTTCATGACAATATGCTGTTTGGAGCTATTCGCAC 658  
118 rIleIleSerCysThrIleIleSerLeuGlyValThrGlnPhePheLysL 135  
659 ACTTATTAGCTGTTCTATTATATATCATTTTGGTGGCTCAAAATTTTCAAGC 708  
135 ysLeuAspIleGlyThrPheAspLeuGlyAspTyrLeuAlaIleGlyAla 151  
709 ACTTAGACATTCACCTTCGGAATTTGGAGATATTTAGCAATTTGGTCGG 758  
152 IlePheAlaAlaThrAspSerValCysThrLeuGlnValLeuAsnGlnAs 168  
759 ATATTTGCTGCAACCGGATTCCTGTTGACATTCGAGGTGCTCAGTCAGGA 808  
168 pGluThrProLeuLeuTyrSerLeuValPheGlyGluGlyValValAsnA 185  
809 TGAGACGGCCCTTACTTACAGTCTCGTGTGTTGGAGAGGGTTCGTCGAATG 858  
185 spAlaThrSerValValPheAsnAlaIleGlnSerPheAspLeuThr 201  
859 ATGCTACATCTGTCGCTCTTTTATGCTATTCATCAAGTTTTCGACATGACT 908  
202 HisLeuAsnHisGluAlaAlaPheHisLeuLeuGlyAsnPheLeuTyrLe 218  
909 AGTTTGTGATCCAAAAATTTGGGCTTCATTTTCATTTGGAACCTTCTTGATT 958  
218 uPheLeuLeuSerThrLeuLeuGlyAlaAlaThrGlyLeuIleSerAlaI 235  
959 ATTTCTCGAGACACTTTTGGGCGTGGGAATTGGACTGCTTTGTGCTT 1008

235 yrValIleLysLysLeuTyrPheGlyArgHisSerThrAspArgGluVal 251  
1009 ATATTATCAAAAAGCTATACCTTTGCGAGGCACTCAACGATCGTGAGTT 1058  
252 AlaLeuMetMetLeuMetAlaTyrLeuSerTyrMetLeuAlaGluLeuPh 268  
1059 GCCCTTATGATGCTCATGCTTACTTGTCTTATATATAATGCGCGAGTTATT 1108  
268 eAspLeuSerGlyIleLeuThrValPhePheCysGlyIleValMetSerH 285  
1109 CTATCTAAGCGGCATACCTTACTGTATTCCTCTGTGGAATTTGTCATGCTC 1158  
285 iSyrThrTrpHisAsnValThrGluSerSerArgIleThrThrLysHis 301  
1159 ATTATACCTGGCACAATGTTACCGAGAGCTCAAGGTCACCTACTAGGCAT 1208  
302 ThrPheAlaThrLeuSerPheLeuAlaGluThrPheIlePheLeuTyrVa 318  
1209 TCCTTTGCAACTCTGTCATTTGTCGAGAGACATTTATCTTCCTCTATGT 1258  
318 lGlyMetAspAlaLeuAspIleAspLysTyrTrpArgSerValSerAspThrp 335  
1259 TGGTATGATGCTTGGATATCGAAAAATGAAATTTGTGAAAAATAGTC 1308  
335 roGlyThrSerIleAlaValSerSerIleLeuMetGlyLeuValMetVal 351  
1309 AGGCACTATCAGTTGTCAGTGAGCTCAATATTTGGTAGGCTTAATCTTAGTA 1358  
352 GlyArgAlaAlaPheValPheProLeuSerPheLeuSerAsnLeuAlaLy 368  
1359 GGCAGAGCTGGTTCGTATCCCTCTGCTTTTATCCAACTTAGCAAA 1408  
368 sLysAsnGlnSerGluLysIleAsnPheAsnMetGlnValValIleTrpT 385  
1409 GAAAACTCTTCGACAGATATCTCTTAGGCAACAAATAATAATTTGGT 1458  
385 rpSerGlyLeuMetArgGlyAlaValSerMetAlaLeuAlaTyrAsnLys 401  
1459 GGGTGGCTTAATGAGAGCGCGCTCTCAATAGCATTGGCTGATAATAAG 1508  
402 PheThrArgAlaGlyHisThrAspValArgGlyAsnAlaIleMetIleTh 418  
1509 TTTACAACCTCGGGGCATCGCTATTCACGAGAACCAATTAATGATTAC 1558  
418 rSerThrIleThrValCysLeuPheSerThrValValPheGlyMetLeuT 435  
1559 AAGTACTGTTACGGTTGTTCTGTTACACAGATTTGTTATTCGGGTTGATGA 1608  
435 hrLysProLeuIleSerTyrLeuLeuPro..... 444  
1609 CGAAGCCTCTGATAAACCCTTCTGCTACCCCGCACAGCAGATGCCAAGC 1658  
445 ...HisGlnAsnAlaThrThrSerMetLeuSerAspAspAsnThrProLy 460  
1659 GGTCAATTCGTCAATGACAACA.....TCCGAACCCAGTAGTCGAA 1699  
460 sSerIleHisIleProLeuLeuAspGln.....AspSerPheIleGluP 475  
1700 GCATCTCAGGTGCGCACTCTCTGGACAACCAACCTGACTCAGAAAGCGATA 1749  
475 roSerGlyAsnHisAsnValProArgProAspSerIleArgGlyPheLeu 491  
1750 TGATAACCGGACCTGAGGTTGCTGACCAACTGCTTGGCGCATGCTGCTA 1799  
492 ThrArgProThrArgThrValHisTyrTyrTrpArgGlnPheAspAspSe 508  
1800 AGGACGCCAACCCACACCGTCACCGCTACTGGCGTAAGTTTGATGATTC 1849  
508 rPheMetArgProValPheGlyArgGlyPheValProPheValProG 525  
1850 GTTTATGCTGCCGCTGTTTGGCGGGCGGATTCGTTCCGTTTGTCCGG 1899  
525 lySerProThrGluArgAsnPro 532



|||||.....  
1900 GCTCACCAGTGTAGCAGAGCCCT 1922

seq\_name: gb\_p11:AB021878

seq\_documentation\_block:  
LOCUS AB021878 2313 bp mRNA PLN 12-AUG-1999  
DEFINITION Oryza sativa OsNHX1 mRNA, complete cds.  
ACCESSION AB021878  
VERSION AB021878.1 GI:5731736  
KEYWORDS  
SOURCE Oryza sativa (cultivar:Nipponbare) cDNA to mRNA.  
ORGANISM Oryza sativa  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
euphyllophytes; Spermatophyta; Magnoliophyta; Liliopsida; Poales;  
Poaceae; Oryza.  
REFERENCE  
1 (sites)  
Fukuda,A., Nakamura,A. and Tanaka,Y.  
Molecular cloning and expression of the Na<sup>+</sup>/H<sup>+</sup> exchanger gene in  
oryza sativa  
Biochim. Biophys. Acta 1446 (1-2), 149-155 (1999)  
2 (bases 1 to 2313)  
Fukuda,A., Tanaka,Y. and Nakamura,A.  
Direct Submission  
Submitted (25-DEC-1998) to the DDBJ/EMBL/GenBank databases.  
Atsunori Fukuda, National Institute of Agrobiological Resources,  
Department of Plant Physiology; Kannondai 2-1-2, Tsukuba, Ibaraki  
305-8602, Japan (E-mail:fukuda@affrc.go.jp, tel:81-298-38-8376,  
Fax:81-298-38-8347)  
Location/Qualifiers  
1 .2313  
/organism="Oryza sativa"  
/cultivar="Nipponbare"  
/db\_xref="taxon:4530"  
297..1904  
/gene="OsNHX1"  
297..1904  
/gene="OsNHX1"  
/note="similar to yeast sodium/proton exchanger"

FEATURES  
source

gene

CDS

/codon\_start=1  
/protein\_id="BAA83337.1"  
/db\_xref="GI:5731737"  
297..1904  
/gene="OsNHX1"  
297..1904  
/gene="OsNHX1"  
/note="similar to yeast sodium/proton exchanger"  
/translation="MGMEVAAARLGAlyTTSYASVSYINLFVALLCACIVLGLLEE  
KRWNESITALLIGCTGVILLMTKGKSSHLFVSEDLFIYLLPLIENAGFOVKK  
NQFENFTILFAGVCTMISFFTSIAAIAIFSRMNLGIDVGDFLAIGAFSATDS  
VCTLOVMDQETPLFLSLVFGGVNDATSIYLFNALQNDLVHIDRAVVLKFLGNFF  
YLFUSSTFLGVFAGLLSAYIIKKLYIGRHSTDEVALMMLMAYLSYMLAELLDDLGIL  
TVFPGIYMSHYTWHNVPESSRVTKHAFATLSIAETFLFYVGMALDIEKWEFAS  
DRPKSIGISLLGLVIGRAAFVFLSFLSNLTKKAPNEKITRQOVVIVWAGLMR  
GAVSTALAYNKTFRSGHTLHGNAIMITSTIVVLFSTVMVGMKPLIRLLLPASGH  
PVTSEPPSPKSLHSPLLTSMOGSDLESTNIVRPSLRMLLTKPTHTVHYWRKFDDA  
LMRPMFGRGVFPSPSPGTSQSGGR"

polyA\_site

BASE COUNT 543 a 538 c 569 g 663 t  
ORIGIN

alignment\_scores:

Quality: 2023.50 Length: 531  
Ratio: 4.121 Gaps: 2  
Percent Similarity: 92.467 Percent Identity: 73.258

alignment\_block:

US-09-271-584A-2 x AB021878 ..

Align seg 1/1 to: AB021878 from: 1 to: 2313

2 LeuAspSerLeuValSerLysLeuProSerLeu...SerThrSerAspHis 17

303 ATGGAGGTGGCGGCGGCTCTGTACAGACCTCCGACTA 352

17 sAlaserValValAlaLeuAsnLeuPheValAlaLeuLeuCysAlaCysI 34  
:|||||.....  
353 CGGTCGGTGGTGTCCATCAACCTGTTTCGTGCGCTGCTCTGGCGCTGCA 402  
34 leValLeuGlyHisLeuLeuGluAsnArgTrpMetAsnGluSerIle 50  
:|||||.....  
403 TCGTCTCGGCGACCTCTCTCGAGGAGAAATCGCTGGTCAATGAGTCCATC 452  
51 ThrAlaLeuLeuIleGlyLeuGlyThrGlyValThrIleLeuLeuIleSe 67  
:|||||.....  
453 ACCGGCTCATATCGGGCTCTGCACCGCGGTGGTGAFTCTTCGTGATGAC 502  
67 rlyGlyLysSerSerHisLeuLeuValPheSerGluAspLeuPhePheI 84  
:|||||.....  
503 CAAGGGAAGAGCTCGCATATTATTCGTCTTCAGTGAGGATCTCTTCTTCA 552  
84 leTyrLeuLeuProProIleIlePheAsnAlaGlyPheGlnValLysLys 100  
:|||||.....  
553 TCTACCTCTCCCTCCGATCATCTTCAATGCAGGTTTTCAGGTAAGAAA 602  
101 LysGlnPhePheArgAsnPheValThrIleMetLeuPheGlyAlaValGI 117  
:|||||.....  
603 AAGCAATTCTTCCGGAATTCATGACGATCACATATTGAGAGCGCTCGG 652  
117 yThrIleIleSerCysThrIleIleSerLeuGlyValThrGlnPhePheL 134  
:|||||.....  
653 GACAATGATATCTTTTTCACAATATCTATTGCTGCCATTCGAATATTC 702  
134 ysLysLeuAspIleGlyThrPheAspLeuGlyAspTyrLeuAlaIleGly 150  
:|||||.....  
703 GCAGAAATGACATTTGGAACGCTGGATGAGAGATTTTCTTGCATTTGGA 752  
151 AlaIlePheAlaAlaThrAspSerValCysThrLeuGlnValLeuAsnGI 167  
:|||||.....  
753 GCCATCTTTCTCGGACAGATTCGTCTGCACATTCAGAGTCCCTCAATCA 802  
167 nAspGluThrProLeuLeuTyrSerLeuValPheGlyGluGlyValValA 184  
:|||||.....  
803 GGATGAGACACCTTTTGTACAGTCTCGTATTTCGGTGAAGGTGTGTGA 852  
184 sAspAlaThrSerValValValPheAsnAlaIleGlnSerPheAspLeu 200  
:|||||.....  
853 ACGATGCTACATCAATTCGTCTTTCACACGCACTACAGAACTTTGATCTT 902  
201 ThrHisLeuAsnHisGluAlaAlaPheHisLeuLeuGlyAsnPheLeuTy 217  
:|||||.....  
903 GTCCACATAGATCGGCTGCTGTTCTGAAATCTCTGGGAACCTCTTTTA 952  
217 rLeuPheLeuLeuSerThrLeuLeuGlyAlaAlaThrGlyLeuIleSera 234  
:|||||.....  
953 TTTATTTTTCGAGCACCTTCTCTGGAGTATTTGCTGCATTCGCTCAGTG 1002  
234 laTyrValIleIleIleIleIleIleIleIleIleIleIleIleIleIle 250  
:|||||.....  
1003 CATACATAATCAAGAGCTATACATTTGGAAGGCATTTCTACTGACCGTGAG 1052  
251 ValAlaLeuMetMetLeuMetAlaTyrLeuSerTyrMetLeuAlaGluLe 267  
:|||||.....  
1053 GTTGCCCTTATGATGCTCATGGCTTACCTTTTCATATATGCTGGCTGAGTT 1102  
267 pHeAspLeuSerGlyIleLeuThrValPhePheCysGlyIleValMetS 284  
:|||||.....  
1103 GCTAGATTTGAGCGGCATTCACCGTATTCCTGCTGTTGTTGTTGTTGTT 1152  
284 erHisTyrThrTrpHisAsnValThrGluSerSerArgIleThrThrLys 300  
:|||||.....  
1153 CACATTACACTGGCATAAACGTCACAGAGAGTTCAAGAGATTACACAAAG 1202  
301 HisThrPheAlaThrLeuSerPheLeuAlaGluThrPheIlePheLeuTy 317  
:|||||.....  
1203 CACGATTTGCAACTCTGCTTCATTCTGAGACTTTTCTCTCTCTCTGTA 1252  
317 rValGlyMetAspAlaLeuAspIleAspLysTrpArgSerValSerAspTr 334





```

LSLKNWPELMKVHVHYLDAMGPFKYPPDVGSLINKLFELLTSLPHVVKDPATPST
RAARLIQCTSTRIATAAEKSVLPHMKGIADTMGYLAKEGTLLRGEHNILSGEAFLVNA
SSAGAOQOEVLAWLEPLLSOOWIQPEMNQNTLSDPMGLLLRVLSLWSPVSPFLPP
EMRAKMTWDTERVSLGEGANPKLSKGVSVVADGSECTKEGOAESESDIRNWLKGI
RCDCNVGLGSLTTIGTEFFKCLDANYVAMALMENLQSEFRHILRFIHTFTYIVKSC
PADMWESVLWGLHPLFIHCOOALSSAWPGLIQEGRKVPDLFGIQSODMKLEWMBE
KLRLDTRIATLFSTMASPGLNTGVPVLEHSHGVGRVDMSTLTLDAFRASNVLLS
LPCLTPNDLHAFEEATATKSPKQKQLMRSLLLIGTGNLKAALAAQKSONVITNVTG
NKVSLHAHDISVFF"
join(<8161..8331,8463..8581,9011..9108,9205..9264,
9340..9386,9475..9669,9764..9810,9892..9958,10061..10161,
10253..10354,10446..10584,10669..10731,10832..10915,
11142..>11489)
/gene="T9J14.2"
8161..11489
/gene="T9J14.2"
/notes="similar to sodium proton exchanger Nhx1 GB:AAD16946
[Arabidopsis thaliana]"
join(8161..8331,8463..8581,9011..9108,9205..9264,
9340..9386,9475..9669,9764..9810,9892..9958,10061..10161,
10253..10354,10446..10584,10669..10731,10832..10915,
11142..11489)
/gene="T9J14.2"
/codon_start=1
/product="putative sodium proton exchanger; 8161-11489"
/protein_id="AAD56988.1"
/db_xref="GI:5932533"
/translation="MTPEASLTSMKLSVSTSDHASVSVSLNLFVALLCACIVIGHLLEE
NRWNESITALLIGLGTVGVILLISRGKNSHLLVSEDLFFIYLLPPIFNAGQVKK
KQFFNFVTIMAFGAIGTVVCTIISLGAIQFFKKLDIGTDFDLGDELAIGAFAATDS
VCTGLAQDETPLLSYVFGVEGVNDATSVLFAIOSFDLTHLNEAPQFGLGNF
YLLFLSTGVATGLISAYIKLYFGRHSTDREVALMMLMAYLSYMAELFSLGSL
TVFFGVIMSHVYTNWNTVSSRTTKHAFATLSLAETFIPLYVGMALDIEKWRFSY
DSPGTSVAVSITLMGLVLMGRNAFVPLSELSNLAKHOSKISIKOOVLTWAGLAR
GAVSALAYNFTSRGHELRGNALMISTITVCLFSTWVFGMLTKPLIRLMPHQKA
TTSSTMSLSDSTPKSIHPLLDGEQDSFELPGSHQDVPKPSLGRFLMRPRTVHY
YWRQDFDAMRPVFGGFGVFPVPGSPTESSHLSKP"
join(<12067..12114,12560..12676,12757..12879,
12957..>13082)
/gene="T9J14.3"
12067..13082
/gene="T9J14.3"
/notes="identical to GB:P11829 [Arabidopsis thaliana]"
join(12067..12114,12560..12676,12757..12879,12957..13082)
/gene="T9J14.3"
/codon_start=1
/product="acyl carrier protein 1 precursor (ACP);
12067-13082"
/protein_id="AAD56989.1"
/db_xref="GI:5932534"
/translation="MATQFSASVSIQTSCLATTRISFOKPALISNHQKTNLSFNLRRS
IPSRRLSYSCAAKQETIEKVAIVKKQLSLTPDKKVAETKPADLGADSLDVEIVMG
LEEFNIOWAEKAKIATVQOAEELIEELINEKK"
complement(join(<13275..13560,13812..13893,13995..14097,
14170..14365,14560..14751,15058..>15342))
/gene="T9J14.4"
complement(13275..15342)
/gene="T9J14.4"
/notes="similar to GB:AA661079 (contains weak similarity to
the SAPB protein (TR:E236624)) [Arabidopsis thaliana]"
complement(join(13483..13560,13812..13893,13995..14097,
14170..14365,14560..14751,15058..15342))
/gene="T9J14.4"
/codon_start=1
/product="unknown protein; 15342-13483"
/protein_id="AAD56990.1"
/db_xref="GI:5932535"
/translation="NRVLSIEAESPFVISRLSPDSTATGGFIGNGVGKCHGLHNLVL
VLAISILFVAYLAYEAKKSLKSLNRRSYIMAYYGFVLWLSLLNLAWCCLQAWBCTPG
KEVINLLTLFTPTSGMLFVLSVAFPLQGNVAGAEALTRTFILISGLIVGLDILLKA
YLFVCGVPPLFDNNEHLHKPKWGLWVHKLLLAGIYGMIFEMVNSKWRRLPCEYID
HSYPYPARAFKYITVMALNGLSLFACALTANGAHFLGLWLYGITSVCYHAFLPLL
YVTFADLFFQEDLNLNENYVSEMKDAGFFDADME"

```

```

mrna
join(16215..16399,16873..16990,17073..17295,17473..17732)
/gene="T9J14.5"
16215..17732
/gene="T9J14.5"
/notes="similar to GB:CAB05547 [Caenorhabditis elegans]"
join(16248..16399,16873..16990,17073..17295,17473..17501)
/gene="T9J14.5"
/codon_start=1
/product="unknown protein; 16248-17501"
/protein_id="AAD56991.1"
/db_xref="GI:5932536"
/translation="MGREVAESCVDTMLMEMVAMYSGRFYANKPELAARRIAGYQV
GHOLSRYTMRERPRFHLAIFICKDFWSEVFKQIDNLKTNHRTGLVQDNKFRW
LSRVSIDPSSNETQDPTPGESKAAQAVSMVLYFPGGIIRGVLSNLGIPCAVSADIS
SLPTCSFVIRKA"
19191..19378
/notes="Emigrant MITE-like element from T517 [17254 to
17779] 526 nt"
<20266..>20949
/gene="T9J14.6"
20266..20949
/gene="T9J14.6"
/notes="predicted by genefinder"
20266..20949
/codon_start=1
/product="hypothetical protein; 20266-20949"
/protein_id="AAD56992.1"
/db_xref="GI:5932537"
/translation="MEKLKADAAEIMICAGQLKGLNLHLMRGMEANTEKSRELDLKEK
ELQILSSDLPQSHAFEAKESEVGDHKLVEECTEELSRKNLLATVKLDSLRVQREL
ELKDNLQVQMAELKRRYSEARHVQRKREMEDEATKKKELSMVTDQIQESGQLEK
KSREVELKDKIEBEKGKELDLVKQVKAWEKLIQLSKLVDDDCCTTELSPRKDOVDS
NNYVVELG"
<22764..>26261
alignment_scores:
Quality: 1924.00 Length: 1110
Ratio: 3.693 Gaps: 22
Percent Similarity: 46.937 Percent Identity: 42.613
alignment_block:
US-09-271-584A-2 x AC009465
Align seg 1/1 to: AC009465 from: 1 to: 93234
1 MetLeuAspSerLeuSerLeuProSerLeuSerThrSerAspHi 17
8167 ATGTTGGCCCTCTTTAACTCTAAAATGCTATCGGTGTCACATCTCGATCA 8216
17 sAlaSerValAlaLeuAsnLeuPheValAlaLeuLeuCysAlaCysI 34
8217 CGCATCTGTCGTTTCACITAACTCTTGTGTCCTTCTATGTGCTTGA 8266
34 leValLeuGlyHisLeuLeuGluGluAsnArgTrpMetAsnGluSerIle 50
8267 TCGTCATGCGCATCTTTTGGAGGAGAAATCGATGATGAACGAATCCATC 8316
51 ThrAlaLeuLeu..... 54
8317 ACTGCTTTATTGATTGTGAAGTTTCCATTATTTTGGCTTGAATGGTAGCTA 8366
54 ..... 54
8367 AAGAGGAGAGCGATAAGTTTACATTTTGGATATTATCATATCAATGTT 8416
55 .....Ile..GlyL 57
8417 TCACCTTTCATTTTCTTTACTGACTTTTAAAGCAATGGGTTGCAGGGGC 8466
57 euGlyThrGlyValThrIleLeuLeuIleSerLysGlyLysSerHis 73
8467 TTGGCACTGGTCTGTCATATGTTGATTAGTAGAGGGAACCAATCACA 8516

```

74 LeuLeuValPheSerGluAspLeuPheIleTyrLeuLeuProIle 90  
|||||  
8517 CTCCTGGCTCTAGTGAAGATCTCTTTATATATCTTTGGCACCAC 8566  
90 eilePheAsnAlaGly ..... 95  
|||||  
8567 AATATTCAATGCAGGGTAAGTTTCTACATCCGGACTGATCATTTGATGAT 8616  
95 ..... 95  
8617 AGATAGTTGCATATTTTGTGCTCTTCTATTTGATACATCTTGTGTCT 8666  
95 ..... 95  
8667 AGTAATATGTTTGGTGGTGTGTACTACAGAGGGGAAAAATCAGTTTA 8716  
95 ..... 95  
8717 TTATCAGAGATCTTACTGGGAAAAATGTTATACTACATTTAGTTTATAAC 8766  
95 ..... 95  
8767 ATATATGCATTATTTGCTGATCTTATGAATATGTTTCAAATATCAGCTTG 8816  
95 ..... 95  
8817 GAAGTACTTGTGATTTGCTACTGCAGAAATAGGAGTGGGTATATGTGAT 8866  
95 ..... 95  
8867 GGCAATTTTGTGATTTATCAATGAAAGTGGTCTTCTATGCTTTGTGCG 8916  
95 ..... 95  
8917 TAAATGCAGACATGGCCAGCTTTCTTTGTTTCACACTATGTTTCTCAG 8966  
96 ..... PheG1 97  
|||||  
8967 ATCTTCTCTGAACCATAATATTATTTGTGCTGTGATTTAGGTTTCA 9016  
97 nValLysLysLysGlnPhePheArgAsnPheValThrIleMetLeuPheG 114  
|||||  
9017 AGTAAAAAGACAGCTTTTTCGAAATTTTGTAACTATATATGCTTTG 9066  
114 lyAlaValGlyThrIleIleSerCysThrIleIleSerLeuGlyVal... 129  
|||||  
9067 GCGCCATTGGGACCGTAGTTTCTTGACCAATAATATCTCTAGGTATCTTA 9116  
129 ..... 129  
9117 CTCCCTCTCCCTGTAAATGATCTGTGTGAAGCTTACTCCTGATTAGTA 9166  
130 ..... ThrGlnP 132  
|||||  
9167 ATTGGGATCAATAATATTTCTGTGATGTTTGAACAGGTGCAATTCAAT 9216  
132 hePheLysLysLeuAspIleGlyThrPheAspLeuGlyAspTyrLeu... 147  
|||||  
9217 TCTTTAAGAAATTAGACATTGGACCTTTGACTTGGCGGATTTCTCTGGT 9266  
147 ..... 147  
9267 AATGTGTTAAGGAGTTCTTTTATCGATTTTGTCTAGAGACTTGAT 9316  
148 ..... AlaIleGlyAlaIlePheAlaAlaThrA 157  
|||||  
9317 TTTTCTTTTGTCTTACTTAATCAGCAATCGGCGCCATATTTGCTGCAACCG 9366  
157 spSerValCysThrLeu.Gln..... 163  
|||||  
9367 ACTCTGTATGCACATACAGGTAATGTGATTCGCCAATTAGCAGAACTAC 9416

163 ..... 163  
9417 ATACAGGCCAGAAATCAAAGTGTAAATTAACAAAAACTCTTGTTCCTTTGG 9466  
164 .....ValLeuAsnGlnAspGluThrProLeuLeuTyrSerLeuVal 177  
|||||  
9467 TGATACAGGTTCTCAATCAAGATGAGACACCTTTGCTTTACAGTCTGTGA 9516  
178 PheGlyGluGlyValValAsnAspAlaThrSerValValValPheAsnAl 194  
|||||  
9517 TTTGGAGAGGGCTTGTGAATGATGCCACATCTGTTGTGCTCTTCAATGC 9566  
194 aileGlnSerPheAspLeuThrHisLeuAsnHisGluAlaAlaPheHisL 211  
|||||  
9567 TATTCAAGAGTTTGTGACCTCACCCACCTTAACCATGAAGCAGCTTTTCAAT 9616  
211 euLeuGlyAsnPheLeuLeuPheLeuLeuSerThrLeuLeuGlyAla 227  
|||||  
9617 TTTCTGGGAACCTTTTATCTGTTTCTTGTGAGCACCGGACTTGGTGTG 9666  
227 ..... 227  
9667 GCAGTAAGTCCATTTTCTCTCTACTGCTCTTTTCGCTGGCTACTACT 9716  
228 .....Ala.Thr 229  
|||||  
9717 TGCAGTTCCTCTGTAACTACAATTTTTTGTATTATCTTGCAGACT 9766  
230 GlyLeuIleSerAlaTyrValIleLysLysLeuTyrPheGly..... 243  
|||||  
9767 GGTCTGATAAGTCTTATGTCATCAAGAACTGTATTTTGAAGGTAGAT 9816  
243 ..... 243  
9817 AATGGCTTTAGCTCATATGTGACTTTTATCTCTATTAGTCTTCTGTTATG 9866  
244 .....ArgHisSerThrAspArgGluValAla 252  
|||||  
9867 TAGTCTAATTTGGGTATCTTTCTTGGGACTCGACTGATCGAGAAAGTTGCC 9916  
253 LeuMetMetLeuMetAlaTyrLeuSerTyrMetLeuAlaGlu..... 266  
|||||  
9917 CTCATGATGTTATGGCTTATCTTTTCATATATGCTTCTGAGGTTAGTTT 9966  
266 ..... 266  
9967 TGTCAAAATTTTATACCACCTGTATAAATTTAGGCATGGAGACAATCAGTTG 10016  
267 .....LeuPhe 268  
|||||  
10017 TATCTTAAGAAATCCTTTATTTGATAATATATTGTTTGGGTTGCAGCTATTC 10066  
269 AspLeuSerGlyIleLeuThrValPhePheCysGlyIleValMetSerHi 285  
|||||  
10067 GCCTTTAGTGTCTATCCTAACTGTATTTTCTGTGGGATTTGATGTCCCA 10116  
285 sTyrThrTrpHisAsnValThrGluSerSerArgIleThrThr..... 299  
|||||  
10117 TTACACTTGGCAAAATGTACCCGAGAGTCAAGAAATTACTACCAAGTAAT 10166  
299 ..... 299  
10167 ATGACCTCCATTTTCTCTTGAACAAATTCAAATGTTTCCCATCTTCAA 10216  
300 .....Lys.HisThrPheAlaIle 305  
|||||  
10217 TACTTACTTACTTTCTATGATTTTACGATGAGACAGCAGCTTTCCTA 10266  
305 hrLeuSerPheLeuAlaGluThrPheIlePheLeuTyrValGlyMetAsp 321  
|||||  
10267 CTTTGTGTTCTCGCTGAGACTTTTATTTTCTCTACGTTGGAATGAT 10316  
322 AlaLeuAspIleAspLysTipArg.Ser..... 330

```
|||||.....|
10317 GCATTGGACATAGAGAAATGAGATTCTGTAGTCACAGGTATGTCCTC 10366
330 .....
10367 AGTAGAAACATAAACTCTCTTTTGTTCATTGGCTATATACATAATCTTC 10416
331 .....ValSerAspThrProGlyThrSerIleAlaVal 341
10417 TTATGTTTCCATTGTTGGTTTGGACAGCCCGGACATCAGTTGCAGT 10466
341 lserSerIleLeuMetGlyLeuValMetValGlyArgAlaAlaPheValp 358
10467 GAGCTCAATCTTAATGGGTCTAGTCATGCTTGGAGAGCAGCTTTGTCT 10516
358 heProLeuSerPheLeuSerAsnLeuAlaLysLysAsnGlnSerGluLys 374
10517 TTCCCTCTTCTTCTTATCAAACTTAGCCAAAGACATCAGAGCGAGAA 10566
375 lleAsnPheAsnMetGln..... 380
10567 ATCAGCATCAAGCAGCAAGTAGCAACAATCTACTTTGTCTATTGAATCA 10616
380 .....
10617 AATTGCTGAGATTGTTGATCTACTTAATCTCTCTCCTTGTCTGT 10666
381 ..ValValIleTrpSerGlyLeuMetArgGlyAlaValSerMetAla 396
10667 AGTTGTGATCTGGTGGCTGTCTTAATGAGAGGTGCTGTATCTATGGCT 10716
397 LeuAlaTyrAsn.Lys..... 401
10717 CTTGCTCAATAAGGTAAATTAAGTAAATGATAGTAACACAGCGAC 10766
401 .....
10767 ATGGTGATTAGTAAANCTACGACATAATTAATCTTTTCTTCTCTG 10816
402 .....PheThrArgAlaGlyHisThrAspValArgGlyAs 413
10817 TGCTCTCTAATGCAGTTTACAAAGATCAGGGCACAGAAATGCGGGGAA 10866
413 nAlaIleMetIleThrSerThrIleThrValCysLeuPheSerThr.... 428
10867 TGCAATCATGATTACCAAGTACAAATACCGTCTGTCTTTTAGCACCATGG 10916
428 .....
10917 TAAGATATCTCCCTCTCAGTTTGGNCTACTAAATGTATTTTCTCTACAG 10966
428 .....
10967 AGATGTTGTCAATGAGGATAAACCAATGCTGAGCGCTGAGACGGATTGC 11016
428 .....
11017 ATTTTCTATATCCGGAATACTGTTCCTCTATTTGATTGTTTATTCCAA 11066
428 .....
11067 TGTTTCTGATGCTAACACTCAACCTTGATACCGGACATTTAATATAT 11116
429 .....Val.ValPheGlyMetLeuThrLysProL 438
11117 GTTTTTTTTTTCCAGGCTTTTCAGGTGTTTGGTATGCTTAACCAAAACCAC 11166
438 eulleSerTyrLeuLeuProHisGlnAsnAla.....ThrThrSer 451
11167 TGATTAGATACCTTAATGCCACCAAAAGCCACCAAGTACCAGAGT 11216
452 MetLeuSerAspAsnThrProLysSerIleHisIleProLeuLeuAs 468
|||||.....|
```

```
11217 ATGTTATCGGAGATAGCACTCCGAAATCAATCCACATTCCTCCTCTCGA 11266
468 p.....GlnAspSerPheIleGluProSerGlyAsnHisAsnValp 482
11267 TGTGTGAACAGCTAGATTCAATTTGAGTTACCTGGGAGCCACCAGGACGTGC 11316
482 roArgProAspSerIleArgGlyPheLeuThrArgProThrArgThrVal 498
11317 CAGACCAACAGCCTTCGAGGTTCTCTCATGCGCCACACGACTGTC 11366
499 HisTyrTyrTrpArgGlnPheAspSerPheMetArgProValPheGl 515
11367 CACTATTACTGGAGACAGTTTGTATGATGCTTCATCGCTCTGTGTTGG 11416
515 yGlyArgGlyPheValProPheValProGlySerProThrGluArgAsnp 532
11417 TGTCTCGGATTCGTTCCCTTGTCTGCTGCTCTCGACTGAGAGNACA 11466
532 roProAspLeuSerLys 537
11467 GCCATGATCTTAGTAAA 11483
seq_name: gb_p12:ATAC011623
seq_documentation_block:
LOCUS ATAC011623 86022 bp DNA PLN 16-NOV-1999
DEFINITION Arabidopsis thaliana chromosome III BAC F24P17 genomic sequence,
complete sequence.
ACCESSION AC011623
VERSION AC011623.3 GI:6437540
KEYWORDS HTG.
SOURCE thale cress.
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
Brassicales; Brassicaceae; Arabidopsids.
REFERENCE 1 (bases 1 to 86022)
AUTHORS Lin,X., Kaul,S., Town,C.D., Benito,M., Creasy,T.H., Haas,B.,
Bonning,C.M., Koo,H., Fujii,C.Y., Utterback,T.R., Barnstead,M.E.,
Bowman,C.L., White,O., Nierman,W.C. and Fraser,C.M.
Arabidopsis thaliana chromosome III BAC F24P17 genomic sequence
Unpublished
REFERENCE 2 (bases 1 to 86022)
AUTHORS Lin,X. and Kaul,S.
Direct Submission
TITLE Submitted (08-OCT-1999) The Institute for Genomic Research, 9712
JOURNAL Medical Center Dr., Rockville, MD 20850, USA, xlin@tigr.org
REFERENCE 3 (bases 1 to 86022)
AUTHORS Lin,X.
Direct Submission
TITLE Submitted (16-NOV-1999) The Institute for Genomic Research, 9712
JOURNAL Medical Center Dr., Rockville, MD 20850, USA
COMMENT On Nov 16, 1999 this sequence version replaced gi:6091707.
Address all correspondence to:
Xiaoying Lin
The Institute for Genomic Research
9712 Medical Center Dr.
Rockville, MD 20850, USA
e-mail: xlin@tigr.org
BAC clone F24P17 is from Arabidopsis chromosome III and is near the
molecular marker mi403.
The orientation of the sequence is from SP6 to T7 end of the BAC
clone.
Genes were identified by a combination of three methods: Gene
prediction programs including GRAIL (available by anonymous ftp
from arthur.epm.ornl.gov), Genefinder (Phil Green, University of
Washington), Gscan (Chris Burge,
http://www.cbs.dtu.dk/netgene/cbsnetgene.html), and NetPlantGene
(http://www.cbs.dtu.dk/netgene/cbsnetgene.html), searches of the
complete sequence against a peptide database and the Arabidopsis
EST database at TIGR (http://www.tigr.org/tdb/at/at.html).
Annotated genes are named to indicate the level of evidence for
their annotation. Genes with similarity to other proteins are named
```

after the database hits. Genes without significant peptide similarity but with EST similarity are named as 'unknown' proteins. Genes without protein or EST similarity, that are predicted by more than two gene prediction programs over most of their length are annotated as 'hypothetical' proteins. Genes encoding tRNAs are predicted as tRNAscan-SE (Sean Eddy, <http://genome.wustl.edu/eddy/tRNAscan-SE/>). Simple repeats are identified by repeatmasker (Arjan Smit, <http://ftp.genome.washington.edu/RM/RepeatMasker.html>). Regions of genomic sequence that are not annotated as genes but have predicted exons by GRAIL are annotated as misc features.

## FEATURES

## source

```

1. .86022
/organism="Arabidopsis thaliana"
/cultivar="Columbia"
/db_xref="taxon:3702"
/chromosome="III"
/map="m1403"
/clone="F24P17"
join(<283. .529,638. .818,916. .1063,1101. .1142,1160. .1376,
1511. .1822)
/gene="F24P17.1"
<283. .1822
/gene="F24P17.1"
/note="similar to pyruvate dehydrogenase kinase
GB:AAC97601, [Arabidopsis thaliana]"
join(<283. .529,638. .818,916. .1063,1101. .1142,1160. .1376,
1511. .1569)
/gene="F24P17.1"
/codon_start=1
/product="putative pyruvate dehydrogenase kinase, 5'
partial"
/protein_id="AAF08568.1"
/db_xref="GI:6437541"
/translation="VRDWLESDRMRAPFEIKSGDEKDEFTOMIKAVKVRHNNVPM
MALGVNQKQKMGNSDLDEHOFDRFLYSRIGRMILGQHVLEHNPPLHTVGYIH
TKMSPEYARNASSEDARSICFREYSGAIEINIGDPSFTFPVPTLHLMYELVKN
LRAVQERVDSDRVPPIRIIVADGIEDVTIKPERSLLHRFPDIIVVSDGGGLARSG
LPRIITYLSTARNPLEDVEDLGIADVPVIMAGVGYGLPISRLRYARFEGDGLIISME
CYGTDAYLHLSRLGDSQEP"
complement(2022. .2051)
/rpt_family="POLY_A"
complement(2053. .2372)
/note="exon predicted by xgrail, quality
excellent_shadowexon"
2057. .2160
/note="exon predicted by xgrail, quality excellent"
2287. .2363
/note="exon predicted by xgrail, quality excellent"
join(<2630. .3712,3813. .4007,4256. .4324,4406. .4675,
4767. .5117,5215. .5492,5579. .>6599)
/gene="F24P17.2"
<2630. .>6599
/gene="F24P17.2"
/note="similar to RNA helicase (DRH1) GB:BAA28347
[Arabidopsis thaliana]; Pfam HMM hits: WW/rsp5/WMP domain
containing proteins, DEAD and DEAH box helicases,
Helicases conserved C-terminal domain"
join(2630. .3712,3813. .4007,4256. .4324,4406. .4675,
4767. .5117,5215. .5492,5579. .>6599)
/gene="F24P17.2"
/codon_start=1
/product="putative RNA helicase"
/protein_id="AAF08584.1"
/db_xref="GI:6437557"
/translation="MATEDTPASAGRPYAPEDTLPQPKGLIDGSTGILYWNPT
NVTQYRPSAPPPIHSATTPKLAQIPVPSGGHOAEQAPVGHVSQGHGFOQPOQ
FPQSHVRPQMQRHQMQPQSQQFPQSQSQSMVPHPHGPSVQTYQPTTQOQOQ
MONHSONPOQLSHOYAHSOQHYMGFRPHOTGLONSHOTPGQPHGQFPQSOEYN
SLAPKREGDEPHGCKTGCQSOPHLPNPSRPSQNTHEEANAASOKTNANLMAQKNG
POANNAVTOFOQGANLIHOOLGPRAPNQDQTMHLHOKSHVSPSQNNTYENHQSRR
GNDSTVNNRMEVVRGAQPLHPAAMPKDIRISGGPTNPADPAMQGTGHGYHAGPAF
PNKSLVRPHFVTSDPVPLHSPVEIYRQHEVTTGTENIPAYITFESSGLPPEILREL

```

```

LSAGFPSPPTQAOQTWPIALQSRDIVAIAKTGSGKTLGYLIPAFILLRHCRNDSRNGP
TVLILATRELATQIODEALRFGRSSRSCTCLYGGAPKGPQLKELERGADIYVATPG
RLNDILEMKMIDFQOVSLLVDEADRLMDMGFEPIRKIYVNEIPPRKQILUMYTATPK
EVRIASDLLLVNPNQVNVGRVDELAANKAITQYEVVVPQMEKERRLEQLILRSQERGSK
VIIFCSTKRLCDHLARSVGRFAGVITHGDKTQGERDVLNQFRSGKSCVLIATDVA
RGLDIKDIRVIVINDFTYGVEDVYVHRIGRTGRAGATGVAFTTEODWKVAPDLIKVL
ESGNOQVPPQVDIRIAMRGGGGPGYSQDRGMVNRFDSSGGGTWDSGGGFGGGRGG
FGAGGFGGREGGFGGREGGFGGRCGRCGRCGRCGRCGRCGRCGRCGRCGRCGRCG
RGFGRCGNNNMESRGFGRCGRCGRCGRCGRCGRCGRCGRCGRCGRCGRCGRCGRCG
RSRSOSYRSRSRSRSRSRSPRHSRDRGHNRSRSYRSRSPVYRPRRDRAPR
VSGFDIKPPVESVNVNLDMAAAAIENVPVPTSLSEROGNGVSEVEAALVRPVVDEEP
"
5843. .5962
/rpt_family="(TGG)n"
complement(7077. .>8090)
/gene="F24P17.3"
complement(7077. .>8090)
/gene="F24P17.3"
/note="similar to hypothetical proteins GB:CAB16818,
GB:AAD55277"
complement(7254. .8090)
/gene="F24P17.3"
/note="unknown protein"
/codon_start=1
/protein_id="AAF08569.1"
/db_xref="GI:6437542"
/translation="MASISLTYLWLVNHPYINFTWIEGETIGSTVFFVSVVSVYL
SATFLURASIDSLPSRILKIPITAVHSLILCLLSLWAVGCTLSITSSHASDDPA
RFLHAICFPYDVKPNPFLFAQVFLSKILEFGDTLILILKTSIORLHFLVYHAT
VVMVYLLWLRTRQSMFPGVLVNSTVHVIMGYVFLCAVGRPKWRLVTDQOIVOFV
FSGLSQWMLRHLFGSGCTGNGWCENAFNASLLALFNSFNHKNVVKPTREDGKK
SD"
8262. .8362
/note="exon predicted by xgrail, quality
marginal_shadowexon"
complement(8375. .8450)
/note="exon predicted by xgrail, quality marginal"
complement(<8719. .>9615)
/gene="F24P17.4"
complement(<8719. .>9615)
/gene="F24P17.4"
/note="similar to hypothetical proteins GB:CAB16818,
GB:AAD55277"
complement(8719. .9615)
/gene="F24P17.4"
/note="unknown protein"
/codon_start=1
/protein_id="AAF08570.1"
/db_xref="GI:6437543"
/translation="MASVSTLYLWLVHHPYIANFTWIEGETIGSTVFFVSVVSVYL
SATFLRYTVDSLPTLGPRIKIPITAVHSLILCLLSLWAVGCTLSLSSDDPKARLF
DAVCFPLDVKPGKPLFFAQVFLSKILEFVDPLLLILNKSQRLSFLHVVHHATYVI
LCYLWLRTRQSMFPGVLVNSTVHVIMGYVFLCAIGSRPKWRLVTFNQVQAFGM
GLGAAMWLPHEHYFGSCAGITVYGVNGFTASLLALFYNFHNSKNYEKTTTPLYKIES
FIHGERWANKAITLFSKND"
9798. .9853
/note="exon predicted by xgrail, quality good_shadowexon"
complement(10146. .11457)
/gene="F24P17.5"
/note="similar to putative protein GB:CAB45783"
/pseudo
complement(12632. .12715)
/note="exon predicted by xgrail, quality
marginal_shadowexon"
join(<12993. .13102,13175. .13257,13357. .13549,13667. .13805,
13900. .14065,14182. .14339,14418. .14612,14700. .14795,
14879. .15013,15093. .15400,15493. .15763,15855. .>16211)
/gene="F24P17.6"
<12993. .>16211
/gene="F24P17.6"
/note="similar to putative Anion exchanger family members:
GB:AAD39673, GB:AAD55295 [Arabidopsis thaliana]"
join(12993. .13102,13175. .13257,13357. .13549,13667. .13805,

```

13900..14065,14182..14339,14418..14612,14700..14795,  
14879..15013,15093..15400,15493..15763,15855..16211)  
/gene="F24P17.6"  
/note="unknown protein"  
/codon\_start=1  
/protein\_id="AAF08571.1"

alignment\_scores:  
Quality: 1492.50 Length: 884  
Ratio: 3.116 Gaps: 22  
Percent Similarity: 54.186 Percent Identity: 42.308

alignment\_block:

US-09-271-584A-2 x ATAC011623 ..

Align seg 1/1 to: ATAC011623 from: 1 to: 86022

2 LeuAspSerLeuValSerLysLeuProSerLeuSerThrSerAspHisAl 18  
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
58793 TTAAGCACAATGCTGGAAAAACAGAGGCTCTTTTCGCTTCTGATCATG 58842  
18 aSerValValAlaLeuAsnLeuPheValAlaLeuLeuCysAlaCysIleV 35  
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
58843 ATCGGTGCTGCCATGAATTTGTTTCGAGCTTTGCTTTGCTTGCAICG 58892  
35 alLeuGlyHisLeuLeuGluAsnArgTrpMetAsnGluSerIleThr 51  
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
58893 TGCCTTGTCACCTGCTTGAGGAGACTCGGTGGATGAACGAGTCAATCACT 58942  
52 AlalaLeuIleLe..... 55  
|||||:|||||  
58943 GCTCTATCATTTGAAGCTTCGTTTTTTTTTTTCCCAAAATCTTAAAC 58992  
56 .....GlyLeuG 58  
|||||  
58993 ACATTCGTTAGAAACATCTAAATTTGGATTTACGCAAAATAGGTTTCGT 59042  
58 lyThrGlyValThrIleLeuLeuIleSerLysGlyLysSerSerHisLeu 74  
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
59043 GTACTGGGATTTGTCATCTTGCTTATAAGTGGAGGCAAAAGCTCAAGGATT 59092  
75 LeuValPheSerGluAspLeuPhePheIleIleThrLeuLeuProIleIle 91  
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
59093 CTTGTGTTTAGTGAAGATCTCTTCTTTATTTATCTTCTCCACCAATTAT 59142  
91 ePheAsnAlaGly..... 95  
|||||:|||||  
59143 ATTCACCGCAGGTAAGTTCTTTTATATAGAAATTCATGAAAAATTGAA 59192  
96 .....PheGlnValLysLys 100  
|||||:|||||  
59193 AGTCGGAATTGAGTAACACATTTTGTGTTACAGGTTTCAGGTTAAGAAG 59242  
101 LysGlnPhePheArgAsnPheValThrIleMetLeuPheGlyAlaValG 117  
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
59243 AAGCAATTTTTCGCACTTCATGACCAATATGTTATTTGGTGTCTATGG 59292  
117 yThrIleIleSerCysThrIleIleSer..... 126  
|||||:|||||:|||||:|||||  
59293 AACTCTCATTTTCATTTGTTATCATCTCATTTGGTAAAAAACCAATCCAA 59342  
126 ..... 126  
59343 TCAACTCCACATTTCAATATCATGAGTCTCATATATTTTGTATGATGTTT 59392  
127 .....LeuGlyValThrGlnPhePheLysLysLe 136  
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
59393 CCGGTTTTGTTTTGTTATGCTAGGTGCTAAACATCTTTTCGAGAAAAAT 59442  
136 uAspIleGlyThrPheAspLeuGlyAspTyr..... 146  
|||||:|||||:|||||:|||||  
59443 GAATATCGGTGATCTTACCATTGCGGACTATCTAGGTACCTGAAACAAA 59492

146 ..... 146  
59493 GAAACCCCTTTTCAAGTCTCTTATTCATTAAAGTCTCATACTGTTTAAATCT 59542  
147 .....Leu..AlaIleGlyAlaIlePheAlaAlaThrAspSer 158  
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
59543 GTACTGATGATTTTGCAGCCATTGGAGCAATATCTCTGCTACAGACTCT 59592  
159 ValCysThrLeu.Gln..... 163  
|||||:|||||:|||||:|||||  
59593 GTTTCACCTTCCAGTAGAAGACTTTTCTGAAACATTCAAAGATATT 59642  
163 ..... 163  
59643 ATTTCTACTACATTTTCTAATCAAAAATTAACCTGCACGCTTTGTTAAATC 59692  
164 .....ValLeuAsnGlnAspGluThrProLeuLeuTyrSerLeuValPheG 179  
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
59693 TAGGTGCTTTAATCAAGACGAGACACTCTCTTGTACAGTCTTGTCTTTGG 59742  
179 yGluGlyValValAsnAspAlaThrSerValValValPheAsnAlaIleG 196  
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
59743 AGAAGGTGTAGTGAAGATGCCACATCGTCTGCTTCAATGCAATAC 59792  
196 InSerPheAspLeuThrHisLeuAsnHisGluAlaAlaPheHisLeuLeu 212  
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
59793 AGAGATTCGACCTCAAAATATCAATTCAGCCATAGCTTTGGAGTTTGT 59842  
213 GlyAsnPheLeuTyrLeuPheLeuLeuSerThrLeuLeuGlyAlaAla.. 228  
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
59843 GGAACATTTTACCTCTTTATCTTAAGCACACGACTTGGTGTTCAGT 59892  
228 ..... 228  
59893 AAGTTTCTTCTCGCTCTCAAAATTTCTAAACCGGAAAAATTTCTCAC 59942  
229 .....ThrGlyLeuIleSerAlaTyrValIleLysLysLeu 240  
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
59943 TCGTTTCTTTTAGGCTGGATTCAGTCTTTTGTATCAAGAAGCTC 59992  
241 TyrPheGly..... 243  
|||||:|||||  
59993 TATATAGGAGGTATATTCGTATCTACTCATGGTTAAAAATGTGATTTA 60042  
244 .....A 244  
60043 CCCTTCTTTTAAATCCTTTTGTCTTTGTTTGTGTTTGTGTTTGAACAT 60092  
244 rg..HisSerThrAspArgGluValAlaLeuMetMetLeuMetAlaTyrL 260  
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
60093 CAGGCACCTCTACTGATCGTGAAGTTGCACCTTATGATGCTATTGGCTTACT 60142  
260 euSerTyrMetLeuAla.Glu..... 266  
|||||:|||||:|||||:|||||  
60143 TATCATATATGTTGGCAGAGTAATAAGAAAAATTTGAACGTTTGGTT 60192  
267 .....L 267  
60193 ACATTTTTCACAGAAAAATGAAACATGATACATATATCCTTTGTCAGC 60242  
267 eupheAspLeuSerGlyIleLeuThrValPhePheCysGlyIleValMet 283  
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
60243 TATTCACCTTGAGCTCTATCTTTGACTGTGTTCTCTGCGGGATTGTTATG 60292  
284 SerHisTyrThrTrpHisAsnValThrGluSerSerArgIleThrThry 300  
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
60293 TCTCATATACATGGCAATGTTACAGATAAATCCAAAGTCACTACAAA 60342  
300 s..... 300  
60343 GTAAGTTTCTTTTCTTCCCAATAAATAGAAAAATAACTTACTTAATTT 60392  
301 .....HisThr 302

```
60393 GTTGTGCTTGCATTGACTAAACTGATACGTTTTTCTTGGCGTAGACATCT 60442
|||||
303 PheAlaThrLeuSerPheLeuAlaGluThrPheIlePheLeuTyrValGI 319
|||||
60443 TTTGCTGCAATGCAATTTCTAGCTAGATTTTATCTCTTTACCTTGG 60492
|||||
319 yMetAspAlaLeuAspIleAspLysTrp..ArgSerValSerAsp..... 333
|||||
60493 AATGGACGCTCTCGATATCGAANAATGGACGTTGTACCAACAGGTGAG 60542
333 ..... 333
60543 ACTTAGAGTGTGATAATACATCAATATCTTTGTTTCAFTTTCATAATAT 60592
334 .....ThrProGlyThrSerIleAlaValSerSerIle 344
|||||
60593 TGTTTAACAAATTTTGGCAGCTCTGGTGGTGGTGGTGGTGGTGGTGGTGG 60642
|||||
345 LeuMetGlyLeuValMetValGlyArgAlaAlaPheValPheProLeuSe 361
|||||
60643 CTCTTTGGGCTTATCTTCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 60692
|||||
361 rPheLeuSerAsnLeuAlaLysAsnGlnSerGlyLeuMetArgGlyAlaValSer 378
|||||
60693 CTCTTTGTCCAAATTAACAAGCTCTTCAACGGATGAGAAATAGACTTAA 60742
|||||
378 snMetGlyValValIleTptTrpSerGlyLeuMetArgGlyAlaValSer 394
|||||
60743 AGAACAAGTAACATTTGGTGGGCTGGTGGTGGTGGTGGTGGTGGTGGTGG 60792
|||||
395 MetAlaLeuAlaTyrAsnLys..Phe..... 402
|||||
60793 ATGGCTCTTGGCTTATAACACAGGTTTGTATCCGTGAACACTACAAGAACA 60842
402 ..... 402
60843 GATAAAAAAGTGATCTTTGGGTTTTTAATAAGTCCCATGACTTAATCTCT 60892
403 .....ThrArgAlaGlyHisThrAspValArgG 412
|||||
60893 TCTCTCTGTATCATATGATGTCACAACTTCAGGACACACCAAGGTTCTTG 60942
412 LysAlaAlaIleMetIleThrSerThrIleThrValCysLeuPheSerThr 428
|||||
60943 GGAACGCTATATGATCATCACCAGTACCATCACTGTGTCTTCTTTCAGTACT 60992
428 ..... 428
60993 GTGGTACGTTTCAACACTTTGTTCAAACCGGTTTGTCTTCTTACTGAAGTC 61042
429 .....V 429
61043 AAATCGGGTTTAGCCTATCCACACGTAACCTAATCTCTAACCTTTTGTG 61092
429 al.ValPheGlyMetLeuThrLysProLeuLeuSerTyrLeuLeuPro.. 444
|||||
61093 CAGGTGTTTGGATTGCTTCAACCAACCGTTAGTCAACATATTTGCAGCCCTC 61142
445 ....HisGlnAsnAlaThrThrSerMet.....LeuSerAspAsp 457
|||||
61143 ATCAAAACAGTCTCCACACGCGCTCGACATCACATAAGATCTTCTTCT 61192
457 snThrProLysSerIleHisIleProLeuLeuAspGlnAspSerPheIle 473
|||||
61193 TCCACCATCCGATCTCCATGAGCGGTGCTCAGTACCACAGGCCAGTC 61242
474 GluProSerGlyAsnHisAsnValProArgProAspSerIleArgGlyph 490
|||||
61243 GAATACGACCTCGAACAACATGTT.....AGCTTCAGAAATGTT 61280
490 eLeuThrArgProThrArgThrValHisTyrTyrTrpArgGlnPheAsp 507
|||||
```

```
61281 CTGGAAATCTCGTCCAGGGCATTTCATCATTTACTTGGAGAAATTCGATA 61330
507 spSerPheMetArgProValPheGlyArgGlyPheValPheValPheVal 523
|||||
61331 ACGCAGTTATGCGTGCATATTTGTCGCCGAGCGGTTTCCACCATGATT 61380
|||||
524 ProGlySerProThrGluArgAsnProProAspLeuSerLys 537
|||||
61381 CCAGTTCACCCATTGAGAAATAGTTCGCCCAATGGAGTGAA 61422
seq_name: gb_p11:AB033990

seq_documentation_block: 7545 bp DNA PLN 06-OCT-2000
LOCUS AB033990
DEFINITION Ipomoea nil Purple gene for Na+/H+ exchanger, complete cds.
ACCESSION AB033990
VERSION AB033990.1 GI:10716130
KEYWORDS
SOURCE
Ipomoea nil (strain:KKZSK-2) foliage DNA, clone_lib:genomic
library.
ORGANISM
Ipomoea nil
Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
Magnoliophyta; eudicotyledons; core eudicots; Asteridae; euasterids
1; Solanales; Convolvulaceae; Ipomoea.
REFERENCE
1 (sites)
Fukada-Tanaka,S., Inagaki,Y., Yamaguchi,T., Saito,N. and Iida,S.
Colour-enhancing protein in blue petals
Nature 407, 581 (2000)
REFERENCE
2 (bases 1 to 7545)
Fukada-Tanaka,S., Inagaki,Y., Yamaguchi,T. and Iida,S.
Direct Submission
Submitted (26-Oct-1999) to the DDBJ/EMBL/GenBank databases.
Yoshihige Inagaki, National Institute for Basic Biology, Division
of Gene Expression and Regulation 1; Myodaiji-cho, Okazaki, Aichi
444-8585, Japan (E-mail:yinagaki@nibb.ac.jp, Tel:81-564-55-7684,
Fax:81-564-55-7685)
FEATURES
Location/Qualifiers
1..7545
/organism="Ipomoea nil"
/strain="KKZSK-2"
/db_xref="taxon:35883"
/chromosome="5"
/clone_lib="genomic library"
/tissue_type="foliage"
join(1355..1525,2541..2659,3187..3284,3400..3459,
3558..3604,3698..3892,4014..4060,4180..4246,4362..4462,
4562..4663,5003..5141,5219..5281,5409..5492,5595..5930)
/gene="Purple"
join(1355..1525,2541..2659,3187..3284,3400..3459,
3558..3604,3698..3892,4014..4060,4180..4246,4362..4462,
4562..4663,5003..5141,5219..5281,5409..5492,5595..5930)
/codon_start=1
/product="Na+/H+ exchanger"
/protein_id="BAB16381.1"
/db_xref="GI:10716131"
/translation="NAFGLSSLLQNSDLFTSDHASVSVSMNLFVALLCACTVILGHLEE
ARWVESITALLIGTGVVILLGGKSHLLVFSEDLFFIYLLPPIIFNAGFQVK
KQFFVNFMTIMLFGAIGTILSCISIFGAVKFKHLDIDFLDFGDIYLAIGAFAATDS
VCTLOVLSODETPLLISLVFGVGVNDATSVLFNAIQSDMTSPDKIGLHFTGNFL
YLFSSFLGVIGILLCAYITKLYFGRHSDREVALMMLMSYLSYIMAEFLYLSGIL
TVFFCGIVMSHTVHNVTSESRVTRHSFATLSFVAETFFLYVGMADALDEKRFVK
NSOGLSVAVSSILYLGRAAFVPLSFLSNLAKNSDKISFRQIIILWAGLMR
GAVSTALAYNRFTTSGHTSLHENAIMITSVTLFSTVFGLMTPLINLLPPHKQ
MPSGHSMTTSEPSPKHFTVPLDNDQPSDESMTITGPPEARPTALRMLLRTPTTHVH
RYWRKDDSFNRVFGGRGFVPFVAGSPVEQSPR"
polyA_signal
6138..6143
polyA_signal
6173..6178
polyA_signal
6184..6189
BASE COUNT 2268 a 1241 c 1325 g 2711 t
ORIGIN
```

```
alignment_scores:
  Quality: 1266.50      Length: 1528
  Ratio: 2.601          Gaps: 26
  Percent Similarity: 31.872  Percent Identity: 26.243

alignment_block:
US-09-271-584A-2 x AB033990 ..
Align seg 1/1 to: AB033990 from: 1 to: 7545

2 LeuAspSerLeuValSerLysLeuProSerLeuSerThrSerAspHisAl 18
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1367 TTGCTCTCTTGGCTCCAAAT...TCGAAATTGTTACGCTCGATCATGC 1413

18 aSerValValAlaLeuAsnLeuPheValAlaLeuLeuCysAlaCysIleV 35
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1414 TTCGGTTGTGTCGATGAACCTCTTTGTGGGGTGTGCTTTGGCGCATGCATG 1463

35 alLeuGlyHisLeuLeuGluGluAsnArgTrpMetAsnGluSerIleThr 51
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1464 TTCFTGGCCATCTACTCGAGAGAATCGCTGGGTGAACGAATCCATCTACT 1513

52 AlaLeuLeuIle..... 55
|||||:|||||
1514 GCCCTTATAATTGTGAGATTTCAGCTTACTATTCTTCATATTGGTTGTA 1563

55 ..... 55
1564 TGTCCCATCTTTAATCAAAACAATGTTGTGTGTGTAATGTTTTTTTT 1613

55 ..... 55
1614 TTTTTTTTTTTTTTTTGGGAAAGAAATAAATATACATCCTTTTTCTG 1663

55 ..... 55
1664 TCCCTATTGGAGTTCAACACCTATGACCTCCCAATTAAAGAGAGTTACAATG 1713

55 ..... 55
1714 ATGCGGATCGGCTCAATACCACAGGCTTTGACAAAATATAAATCTGTTA 1763

55 ..... 55
1764 GATCATAGCCGGGCCGACATGAGGGTGGCGGATATGGCACCTGGGCCAC 1813

55 ..... 55
1814 CAACGTTGGAGGGCCCTAGCTTGGTTGTTTTTTTTTTTATTATTATT 1863

55 ..... 55
1864 TATTTATTTTATATTTTGTATGATGAGCTTAAAAAATGTTCCAAATGCAA 1913

55 ..... 55
1914 GATTTGAATTCATGATTTCAAAATGTTTATTAATTTTAAATCCACTAA 1963

55 ..... 55
1964 CAATTACAACAATGATATATCTACTTATCTCAITTAATACATACAGAGTATA 2013

55 ..... 55
2014 TATTAGTGAACCTTGAATTAATTTAATTAATAATAAATCTCATATAGGCCTAA 2063

55 ..... 55
2064 AGTTTGAATGGCCTACTCCAAAGATTAACTAAAAAATATATATTTTTT 2113

55 ..... 55
2114 AAAATAAATTATATGATATACAAAAAATTCAAAATTTATCATTTAATTATTTA 2163
```

```
55 ..... 55
2164 TAATTAACACATACAAATCCTTTAACTTTCATATATTAAGAGCCCTTGATT 2213

55 ..... 55
2214 TTTTTTTTTTGAAGTAAAGACCCCTTGTGTTTTATTATTTGCCCGGAATCTATA 2263

55 ..... 55
2264 AAATGATAGGATCGCCCTGGATCATAGAGTAAACCTCCAAAAGCCCTCTA 2313

55 ..... 55
2314 GCATGTGAGTCTGGTTAGGAATCATTTGGTTTAATATTCATTTCTCTCAA 2363

55 ..... 55
2364 AVATACAAAAAAGTCAGTGTGTGAGTGACATGTAAGATAAAGGTATA 2413

55 ..... 55
2414 TGGTTAGCTGATTAATAGCATTTTGTAGAATTTAATTAATATAGTTTG 2463

55 ..... 55
2464 TTATTTTCATCATATTATTGTTATTTTAACTTTTATGTTCTCTTCACAAT 2513

56 ..... Gly. LeuGlyThrGlyValThrI1 63
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
2514 TGTATTTTATTATGATGATGAGGAGGTTTGTGCACGGAGTTGTAAT 2563

63 eLeuLeuIleSerLysGlyLysSerSerHisLeuLeuValPheSerGluA 80
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
2564 TTTGCTCTTAGCGGAGAAAGATTACATCTCTCTCTTTTAGCGAAG 2613

80 sLeuPhePheIleTyrlLeuLeuProIleIlePheAsnAlaGly... 95
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
2614 ATCTTTCTTTATATATCTCTGCGCACCTATATATTCATGCG.GGGTA 2662

95 ..... 95
2663 AATCTTATGATTTTTTTTTTTTTTTTTTTTTTTTACCTCTATATGTTGTTT 2712

95 ..... 95
2713 ACTGGTGAATTTATATATCTCTCTCTGATATAAAATGCTGTGTGGGGGAAA 2762

95 ..... 95
2763 TTAACAGTTATTACCTTCATCTCTCTCTCTTTTGACATTGAATTAGTCTT 2812

95 ..... 95
2813 GTCTCATATAATTAATAAATCCAATCATTTATGTCACACACACGCGTATA 2862

95 ..... 95
2863 GACCATAGTATCAATAAATCTATACACTACACAAATAATGTAATTTAGT 2912

95 ..... 95
2913 ATATTAAAAATGTACATTGTATTAGTAACATAGTTTGTGAAGCATTTT 2962

95 ..... 95
2963 TCGTGAAAAATTCATATATGTTTTTATGGGTGGTATCTTTGTTTCACAATCAA 3012

95 ..... 95
3013 TAACTAATTCAGGAAATACATGTAATGAGCATTCCTTTTATTGAG 3062
```



95 ..... 95  
3063 GTTGTAGAGTAGTTATTGATTCTTAACCTTGAATAATTTTGTGGTAATC 3112  
95 ..... 95  
3113 TCGTGGTTCAGTGTCTAATCTGCCAAATAAGTGTCTTTTTTTTTTTTTTT 3162  
96 ..... PheGlnValLysLysGlnPheP 104  
3163 TTATAATGATTACCTATACACGGTTTCAAGTGAAANAAGACAGCTTT 3212  
104 heAraSnpPheValThrIleMetLeuPheGlyAlaValGlyThrIle 120  
he .....  
3213 TCGTGAACCTTCATGACAATTAATGCTGTTGAGCTATTGCCACACTATT 3262  
121 SerCysThrIleIleSer. Leu. 127  
SerCysThrIleIleSer. Leu. 127  
3263 ACTGTGCTATTATATCAATTTGGTACTTAATCAAAACCTCGACATCTTAT 3312  
127 ..... 127  
3313 TACCAGCTAATGCCCTTCTTTGTCACAAATAAATTTATGCTGCGATTGATG 3362  
128 ..... GlyValThrGlnPhe 132  
3363 TATTATCAACTCATATTCCGTATCTATAGATTAGGTGCGGTCAAAAT 3412  
132 ePheLysLysLeuAspIleGlyThrPheAspLeuGlyAspTyr. .... 146  
ePheLysLysLeuAspIleGlyThrPheAspLeuGlyAspTyr. .... 146  
3413 TTTCAGCACTTATGACATTTGACTTCTGATTTTGGAGATTATTTAGGTA 3462  
146 ..... 146  
3463 TCATGTATATTCGCTCTCTTTACATACACACAGTATCTTCATGCTTCT 3512  
147 ..... Leu. AlaIle 149  
3513 TGATTCTAAACAATTTTTTTTTTCTCTTTTATTATTCACAATTCAGCAAT 3562  
150 GlyAlaIlePheAlaAlaThrAspSerValCysThrLeuGlnValLeuAs 166  
3563 GGTGGCATATTTGCTGCAACCGATTCTGTTTGCACATTGAGGTACTGAA 3612  
166 n ..... 166  
3613 TATTATGATTATGAAGTACGCTATTCATTCGGTGAAGGCACATATTAA 3662  
167 ..... GlnAsp 168  
3663 TTATAGCTAAAAATAAATTTTATTTATGCGACAGGTGCTCAGTCAGAT 3712  
169 GluThrProLeuLeuTyrSerLeuValPheGlyGluGlyValValAsnAs 185  
3713 GAGAGCCCTACTTTACAGTCTCGTGTGTTGGAGAAGGGTGGTCAATGA 3762  
185 pAlaThrSerValValPheAsnAlaIleGlnSerPheAspLeuThrH 202  
3763 TGCTACATCTGTGGCTCTTTTATGCTATTCAAGTTTTCACATGACTA 3812  
202 isLeuAsnHisGluAlaAlaPheHisLeuLeuGlyAsnPheLeuTyrLeu 218  
isLeuAsnHisGluAlaAlaPheHisLeuLeuGlyAsnPheLeuTyrLeu 218  
3813 GTTTTGATCCAAAAATGGGCTTCATTTTCATTTGGAAACTCTTGTATTTA 3862  
219 pheLeuLeuSerThrLeuLeuGly ..... 226  
pheLeuLeuSerThrLeuLeuGly ..... 226  
3863 TTCTCTCGAGCACATTTTGGCGCTGGAGTAAGTTCTTCTCCCGACTA 3912  
226 ..... 226  
3913 GCATATCTCACCCCGTGTGTTGTCTCTATCCCGGTGTTTCTGTGTC 3962  
227 ..... AlaAla 228

[illegible]

|      |  |      |
|------|--|------|
| 4863 | CATGAGTTAGACTAACTAAATTGATTGGATGCAGTCCTTTTAGTGTGGTTTA | 4912 |
| 335  | .....  | 335  |
| 4913 | TTATAAGTTATAACAATTTTAAAGATAAAACAATAAACTATGTTTCACG    | 4962 |
| 336  | .....GlyThr  | 337  |
| 4963 | GAACCTTATAAATCTCTTGTTATTATTATTCTTTCAGTCAGGGACTA      | 5012 |
| 338  | SerIleAlaValSerSerIleLeuMetGlyLeuValMetValGlyArgAl   | 354  |
| 5013 | TCAGTTGCAGTGAGCTCAATATTGGTAGCCCTAATCTTAGTAGCAGAGC    | 5062 |
| 354  | aAlaPheValPheProLeuSerPheLeuSerAsnLeuAlaLysLysAsnG   | 371  |
| 5063 | TGCGTTCGTATTCCCTCTGCTGTTTATCCAACTTAGCAAGAACT         | 5112 |
| 371  | InSerGluLysIleAsnPheAsnMetGlnVal..Val.....           | 382  |
| 5113 | CTTCGACAAGATATCCTTTAGGCAACAAGTAAGAGTCTCTACAAAT       | 5162 |
| 382  | .....  | 382  |
| 5163 | TCTCATTTTATTATTATCTTTACGAGTAATATACATCTCTTACTCTGTC    | 5212 |
| 383  | .....IleTrpTrpSerGlyLeuMetArgGlyAlaValSerMe          | 395  |
| 5213 | TCTTAGATAAATAATTTGGTGGCTGCCTAATGAGAGCGCGTCTCAAT      | 5262 |
| 395  | tAlaLeuAlaTrAsn..Lys.....                            | 401  |
| 5263 | AGCACTTGCGTATAATVAGGTAAGTATAATAGTCACACAAATGTTA       | 5312 |
| 401  | .....  | 401  |
| 5313 | CAATGTCACCTTAGAATTGAAAGACCGGTGGTGGTCTTCTCCATTTCC     | 5362 |
| 402  | .....PheT  | 403  |
| 5363 | CGATTTCCCTGTGATTCTGTGATTAGATGTTATTGACGCGGTGCAGTTTA   | 5412 |
| 403  | hArgAlaGlyHisThrAspValArgGlyAsnAlaIleMetIleThrSer    | 419  |
| 5413 | CACCTCGGGCAGTAGCTATTCACGAGACGCAATAATGATTACAAGT       | 5462 |
| 420  | ThrIleThrValCysLeuPheSerThrVal.....                  | 429  |
| 5463 | ACTGTTACGGTGTCTGTTTCAGCAGAGTTGTAAGGACCTACATACAT      | 5512 |
| 429  | .....  | 429  |
| 5513 | TTTCGAGGCTTTTCTTTGATTCTCGCCAAATGAATCTACGGTTGG        | 5562 |
| 430  | .....ValPheGlyMetLeuThr                              | 435  |
| 5563 | TTGGTTGGTTGTGTTGGTGATGTTTATTAGTATTTCGGTTGATGACG      | 5612 |
| 436  | LysProLeuIleSerTyrLeuLeuPro.....                     | 444  |
| 5613 | AAGCCTCTGATAACCTTCTGCTACCCCGCACCAAGCAGATGCCAGCGG     | 5662 |
| 445  | HisGlnAsnAlaThrThrSerMetLeuSerAspAsnThrProLys        | 461  |
| 5663 | TCATTCTGCTAATGACACA.....TCCGAACCCAGTAGTCCGAAGC       | 5703 |
| 461  | erIleHisIleProLeuLeuAspGln.....AspSerPheIleGluPro    | 475  |
| 5704 | ACTTCACGGTGCCACTCCTCGACACCAACCACTGACTCAGAAACGATATG   | 5753 |
| 476  | SerGlyAsnHisAsnValProArgProAspSerIleArgGlyPheLeuTh   | 492  |
| 5754 | ATAACCGGACTGAGGTGCTCGACCAACTGCCTTGCCTGCTGCTAG        | 5803 |

seq\_name: gb\_p11:AB015479

seq\_documentation\_block: 80675 bp DNA PLN 09-AUG-2000

LOCUS AB015479 Arabidopsis thaliana genomic DNA, chromosome 5, pl clone:MTE17.

DEFINITION AB015479

ACCESSION AB015479

VERSION AB015479.1 GI:3241927

KEYWORDS Arabidopsis thaliana (strain:Columbia) DNA, clone\_lib:Mitsui P1 clone:MTE17.

SOURCE Arabidopsis thaliana

ORGANISM Arabidopsis thaliana

Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

REFERENCE 1 (sites)

AUTHORS Nakamura,Y., Sato,S., Asamizu,E., Kaneko,T., Kotani,H., Miyajima,N. and Tabata,S.

TITLE Structural analysis of Arabidopsis thaliana chromosome 5. VII. Sequence features of the regions of 1,013,767 bp covered by sixteen physically assigned P1 and TAC clones

JOURNAL DNA Res. 5 (5), 297-308 (1998)

MEDLINE 99087489

REFERENCE 2 (bases 1 to 80675)

AUTHORS Nakamura,Y.

TITLE Direct Submission

JOURNAL Submitted (17-JUN-1998) to the DDBJ/EMBL/GenBank databases.

Yasukazu Nakamura, Kazusa DNA Research Institute, Department of Plant Gene Research; 1532-3, Yana, Kisarazu, Chiba 292-0812, Japan (E-mail:ynakamu@kazusa.or.jp, Tel:81-438-52-3935, Fax:81-438-52-3934)

COMMENT Address for correspondence: kaos@kazusa.or.jp

For the latest information on annotation of this clone, please see http://www.kazusa.or.jp/kaos/cgi-bin/agd\_graph.cgi?c=MTE17

Genes with similarity to proteins in the databases are described in 'product' or 'note' qualifiers. Genes that have no significant protein similarity are described as 'unknown protein'.

The software programs used to predict genes include: Grail (Informatics Group, Oak Ridge National Laboratory, http://compbio.ornl.gov/grail-1.3/), GENSCAN (Chris Burge, MIT, http://CCR-081.mit.edu/GENSCAN.html), NetGene2 (S.M. Hebsgaard, et al., CBS, Technical University of Denmark, http://www.cbs.dtu.dk/services/NetGene2/) and SplicePredictor (Volker Brendel, Stanford University, http://gremlini.zoool.iastate.edu/cgi-bin/sp.cgi), genes encoding tRNAs are predicted by tRNAscan-SE (Sean Eddy, Washington University School of Medicine, St. Louis, http://genome.wustl.edu/eddy/tRNAscan-SE/).

this sequence may not be the entire insert of this clone. It may be shorter because we remove overlaps between neighboring submissions.

The 5' clone is MCOL5 and the 3' clone is MWC10.

FEATURES

Location/Qualifiers

1. .80675

/organism="Arabidopsis thaliana"

/strain="Columbia"

/db\_xref="taxon:3702"

/chromosome="5"

/clone="MTE17"

/clone\_lib="Mitsui P1"

complement(join(636..713,796..964,1048..1151,1258..1347,1437..1505,1645..1778,1867..1957,2059..2139,2265..2390,

CDS



|                     |        |                   |        |
|---------------------|--------|-------------------|--------|
| Ratio:              | 2.529  | Gaps:             | 24     |
| Percent Similarity: | 43.719 | Percent Identity: | 30.762 |

alignment block:

US-09-271-584A-2 x AB015479/rev

Align seq 1/1 to reverse of: AB015479 from: 1 to: 80675

15 SerAspHisAlaSerValValAlaLeuAsnLeuPheValAlaLeuLeuCy 31  
:::| | | | | | | | | | | | | | | | | | | | | |  
48342 GCTGAGCATCTCAGGTGATACCAATCTCAGTGTTCATCGCCATTCTCTG 48293

31 sAlaCysIleValLeuGlyHisLeuLeuGluGluAsnArgTrpMetAsnC 48  
| | | | | : | | | | | : | | | | | : | | | | | : | | | | |  
48292 TCATGTTT AGTTATCGGCCACTTGCTTGAGAGAAATCGATGGGTAAATG 48243

48 luSerIleThrAlaLeuLeu..... 54

|    |       |    |
|----|-------|----|
| 54 | ..... | 54 |
|----|-------|----|

48192 TTTTGTTTTAGCTAAAGTTTGGAGCTTTTCTGCTCTTAAGAACAATTT 48143  
55 ILe G[V]LeuS[V]ThrG[V]Va 61

48142 ATCACTATTTTGTGATTGGTTTGTTCGAAGTAGGGAGCAGCATCAGGAAC, 48093

48092 AGTGATCTTACTTATTAGTAAAGGAAAAAGTTACATATATTTGGTGTG 48043

**78** erGluAspLeuPhePheIleTyrLeuLeuProProfilellePheAsnAla 94  
:  
:::  
:

|              |    |
|--------------|----|
| 95 .Gly..... | 95 |
|--------------|----|

4/99Z GGGTAAGCTTACTATGGTTTGGTTACTATGTTGGTTAAAGGGATTAGA 4/94A

95 ..... 95

47942 TTACACTATGATGGTGATTGACTGAAGATTTTTGGTTGTGTGGCAGG 47893

47892 TTCCAAGTTAAGAAAAAGAAGTTTTTTCACAACCTTTTAACCATCATGTC 47843

112 upheGlyAlaValGlyThrIleLeSerCysThrIleLeSerLeuGlyW I29  
|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::  
47842 CTTTCCTGTGATTGGAGTTTTTCAATCCACACTGCATATCATCGTTTGGA 47793

129 a l t h r G l n P h e P h e L y s L y s L e u A s p I l e G l y T h r P h e A s p L e u G l y . . A 145

|                   |     |
|-------------------|-----|
| 145 spTyrLeu..... | 147 |
|-------------------|-----|

47742 ATTATCTCGGAGTTGTTACCATCTCTATTCTTTGTTGTTTTATAATGTA 47693

47692 ACCAGGGACTTGGTGGCTGTTTCCCAAGTTGGGATTTAAGGGGTTGAGTG 47643

14/ ..... 47642 CTAGAGACTATCTTGGTAAGATTTTCACTGATCTATGTTTCTTTTCATT 47593

|     |       |     |
|-----|-------|-----|
| 147 | ..... | 147 |
|-----|-------|-----|

148 .....AlaIleGlyAlaIlePheAlaAlaThrAspSerValCy 160

47542 TGTTTTCGATTCAGCCATAGGACGATTTTCTCATCACTGATCTGTTG 47493  
|||||.....  
160 sThrLeuGlnVal.....LeuAsn 166  
|||||.....  
47492 CACTCTACAGGTTCTTGATTAGGAAGATCTTGTCTTACTACTTCTCT 47443  
|||||.....  
165 .....LeuAsn 166  
|||||.....  
47442 ATTAAGAGTTTTTCTTACATTCGCATCTTAACCTTTTACAGATTCCTCAT 47393  
|||||.....  
167 GlnAspGluThrProLeuLeuTyrSerLeuValPheGlyGluGlyValVa 183  
|||||.....  
47392 CAAGATGAACACCATTCCTATACAGCTTAGTCTTTGGAGAAGCAGTGT 47343  
|||||.....  
183 IAsnAspAlaThrSerValValValPheAsnAlaIleGlnSerPheAspL 200  
|||||.....  
47342 CAATGATCAACCTCAGTTGTACTGTTCACGCGCTGCAAAGATTCAT 47293  
|||||.....  
200 euThrHisLeuAsnHisGluAlaAlaPheHisLeuLeuGlyAsnPheLeu 216  
|||||.....  
47292 TTGAAGCCTAACCGTTGGAGCGCTGCAAGTATTATGGGAACCTTTTG 47243  
|||||.....  
217 TyrLeuPheLeuLeuSerThrLeuLeuGly..... 226  
|||||.....  
47242 TACCTCTTCTCAACAAGCACACTTCTCGGAATTGGTGTAAAGCTATTGAGT 47193  
|||||.....  
226 ..... 226  
|||||.....  
47192 TGGCATAGTTTTTTTACTCTGTTTGAAGCATATCATAGTAATCTTAGAAC 47143  
|||||.....  
227 .....AlaAla..ThrGlyLeuIleSerAlatyrValIleLysLysL 240  
|||||.....  
47142 TCTGAATCTCTTGCAGGTGGGCTTAATAACATCTTTTCTTCTTAAACCT 47093  
|||||.....  
240 euTyrPheGly..... 243  
|||||.....  
47092 TGTATTTTGGAAAGGTAAAGTACACAACTTGCAATTTTTCATCTCAA 47043  
|||||.....  
243 ..... 243  
|||||.....  
47042 GCTTAAGTTTTTTTTTTTATAGTATAAGATGTCCTTTATATTGATCTTG 46993  
|||||.....  
244 .....ArgHisSerThrAspArgGluValAlaLeuMetM 255  
|||||.....  
46992 TTATATGGTTGCTGCAGACAGACATCTACTACACGCGAAGCTCGCCATCAG 46943  
|||||.....  
255 eLeuMetAlaTyrLeuSerTyrMetLeuAlaGlu..... 266  
|||||.....  
46942 TTCTAATGGCTTACCTTTTCATATATGTTGGCTGAGGTAAATACACTGGTCA 46893  
|||||.....  
266 ..... 266  
|||||.....  
46892 TGAACGCTACTTTTTAGTTTTTTCATTTGGAAATTGAAGTGTGCTGCTTCTCT 46843  
|||||.....  
267 .....LeuPheAspLeuSerGlyI 273  
|||||.....  
46842 GTTTTAACTCAATATAGTTGTTCTCTGCAGCTCTTCTCATTAAGTGGAA 46793  
|||||.....  
273 leuLeuThrValPhePheCysGlyIleValMetSerHisTyrThrTrpHis 289  
|||||.....  
46792 TTCCTACTGTTTCTCTCTGGTGGTTTTAATGTCGCATTATCGCATCATAT 46743  
|||||.....  
290 AsnValThrGluSerSerArgIleThrThr..... 299  
|||||.....  
46742 AACGTCACAGAGAGCTCAAGAATCACTTCCAGGCAATGTATCAGACAAAT 46693  
|||||.....  
299 ..... 299  
|||||.....  
46692 GTTATTCGCTGAAGTCTCAATTTGCAGGAATTCGTTTTTGTCTGACTGTG 46643  
|||||.....  
300 .....LysHisThrPheAlaThrLeuSerPheLeuAlaGlu 311  
|||||.....  
|||||.....

46642 CATCTATATTTTGCAGCATGATTTGCAATGTTGTCCTTTTATTCGGGAG 46593  
312 ThrPheIlePheLeuTyrValGlyMetAspAlaLeuAspIleAspLysTr 328  
|||||  
46592 ACATTCATATTTCTGTATGTTGGACAGATGCTCTGTATTTTACAAAGTG 46543  
328 pArgSerValSer..... 332  
|||||  
46542 GAAGACAAGCAGCTTAAGGTATAAAGTCGTATCCATTTATCAGCTGGTAG 46493  
332 ..... 332  
46492 AGCGTTTACAGTGTTTGTGCAAAACAATACACAGCATTTTGTATATGAACC 46443  
333 ..... Asp.ThrProGlyThrS 338  
|||||  
46442 TCGTGTATTTCTTAATGCATACCTAGATCAAGATTTATCCCAAGACAT 46393  
338 er..... 338  
46392 CCTCTATGCTTCTTCTCTCTTTGACACGAAGTTAATCGCAATTAACTGT 46343  
338 ..... IleAlav 341  
46292 AGACTTGTAAACATTAATCATCTTCTTGACGTTTGGGGGTACTCTGGGTG 46243  
341 alSerSerIleLeuMetGlyLeuValMetValGlyArgAlaAlaPheVal 357  
|||||  
46242 TCTCCGGGTGCATAACCGCATTTAGTATTGCTGGACGACGACATTTGTC 46193  
358 PheProLeuSerPheLeuSerAsnLeuAlaLysLysAsn..... GlnSe 372  
|||||  
46192 TTTTCACCTCTCGGCTTCAACAATTTTCATGAACAGGCACACTGAAAGAAA 46143  
372 rGluLysIleAsnPhe.AsnMetGln..... 380  
46142 CGAGTCTATCACATTTAAGCATCAGGTTTACTCTTGTCTTTTACAGTTA 46093  
380 ..... 380  
46092 CATTATCAGAAGCACTAAATAAAAAAATTAATATGTCGCCAGATGGA 46043  
381 ..... ValValIleT 384  
46042 AGTTTCTGTTATATAAATCTCTCTTCTGTATACATTTATAGTGATCATTT 45993  
384 rpTrpSerGlyLeuMetArgGlyAlaValSerMetAlaLeuAlaTyrAsn 400  
45992 GGTGGGCAGGCTAATGCGAGTGCTGCTCAATTGCTGCTGCTTCAAG 45943  
400 ..... 400  
45942 CAGGTTTTTCTCTTGTGCATTTATATCTCTAACTCTCTATTGTATTA 45893  
400 ..... 400  
45892 CTTTACATTTGAATTAGTAACATTTCCATTTATATCTCATTTTGGTATATC 45843  
401 ..... LysPheThrArgAlaGlyHisThrAspValArgGlyAsn 413  
45842 TTTTCATGGCCAAAGTTCATACCTACTCCGGTGTATCATTTGGATCCTGTGAAT 45793  
414 AlalleMetIleThrSerThrIleThrValCysLeuPheSerThrVal... 429  
45792 GCTGCCATGGTCACCAACACCACTATCGTTGCTCTCTTCTTACTACACTGGT 45743  
429 ..... 429  
45742 AAACAATACTACCTTAGTATATATGGCCCTACTCGTTCTGCTAGTATTTCT 45693

430 ..... ValPhe 431  
45692 GTGTTAAGATCAAAAGATGACTTTTGTGTCATGTAACAACAGGTCTTT 45643  
432 GlyMetLeuThrLysProLeuIleSerTyrLeuLeuPro..... 444  
45642 GGTTCCTCACAACACCATTTGTAATTAICTCTTCTCAAGATGCAAG 45593  
445 HisGlnAsnAlaThrSerMetLeuSerAspAspAsnThrProLys. 460  
45592 TCACAACACCGGAATAGAGTAACACGACTGAGCCAGGTTCTCCGAAG 45543  
461 SerIleHisIleProLeuLeuAspGlnAspSerPheIleGluProSer 476  
45542 AAGATGCGACACTTCTCTCTCTT.....TCCTTTGACGAGTCTGCT 45502  
477 GlyAsnHisAsnValProArgProAspSerIleArgGlyPheLeuThrAr 493  
45501 TCCACCACTTCATAGACTAAGATAGTATTTCCTCTCTGATGGGAACA 45452  
493 gProThrArgThrValHisTyrTyrArgGlnPheAspAspSerPheM 510  
45451 ACCTGTGTACACCATCCACCGCTACTGGAGAAAGTTTGACGACACATACA 45402  
510 etArgProValPheGlyGlyArgGlyPheValPropheValProGlySer 526  
45401 TGAGGCTATCTTTCGGTGA..... 45382  
527 ProThrGluArgAsnProProAsp 534  
45381 CCTCGTCGAGAAAACCAACAGAA 45358  
seq\_name: gb\_in2:AF142676  
seq\_documentation\_block: 2405 bp mRNA INV 27-MAY-1999  
LOCUS AF142676 Drosophila melanogaster sodium-hydrogen exchanger NHE1 (NHE1) mRNA,  
DEFINITION complete cds.  
ACCESSION AF142676  
VERSION AF142676.1 GI:4894990  
KEYWORDS fruit fly.  
SOURCE Drosophila melanogaster  
ORGANISM Drosophila melanogaster  
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;  
Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.  
REFERENCE 1 (bases 1 to 2405)  
AUTHORS Dow,J.A.T.  
TITLE An insect member of the Na+/H+ exchanger family  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 2405)  
AUTHORS Dow,J.A.T.  
TITLE Direct Submission  
JOURNAL Submitted (14-APR-1999) IBLs-Molecular Genetics, University of  
Glasgow, Anderson College, Glasgow, Scotland G11 6NU, UK  
FEATURES  
source  
1..2405  
/organism="Drosophila melanogaster"  
/db\_xref="taxon:7227"  
1..2405  
/gene="NHE1"  
32..1717  
/gene="NHE1"  
/codon\_start=1  
/product="sodium-hydrogen exchanger NHE1"  
/protein\_id="AAD32689.1"  
/db\_xref="GI:4894991"  
/translation="MRVWVAYSAALLLVHAGPESIGQGVPOSKYSSNTTTTDDSS  
SIHTVSDVFNPSLQNTVPSISAGSNASTTKRGNASTLVTDPLIDSHAVEQEHNSL  
SLFEFVICVIMLGILLIHSMLQTFQYLVSTIVVFLGAFGLSLNVMGONGWKRE  
VFSPMGFLLVLLPTIFESYGNLHKGNFQNGISLVFAIFGTTISALVIGAGIYLLG  
LGEVAFRLSFSESAFGLSISAVDPVATVATFALHDVDPILNMLVFESILNDAISIV



```

mRNA, partial cds.
ACCESSION AF279671
VERSION AF279671.1 GI:9652169
KEYWORDS common ice plant.
SOURCE Mesembryanthemum crystallinum
ORGANISM Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
Magnoliophyta; eudicotyledons; core eudicots; Caryophyllales;
Caryophyllales; Aizoaceae; Mesembryanthemum.
REFERENCE 1 (bases 1 to 421)
AUTHORS Chauhan,S., Forsthoefel,N. and Bohnert,H.J.
TITLE Na+/myo-inositol symporters and Na+/H+-antiporters in plant sodium
transport
JOURNAL unpublished
REFERENCE 2 (bases 1 to 421)
AUTHORS Chauhan,S., Forsthoefel,N. and Bohnert,H.J.
TITLE Direct Submission
JOURNAL Submitted (19-JUN-2000) Biochemistry, University of Arizona, 1041
E. Lowell St, Tucson, AZ 85721, USA
FEATURES
Source
location/Qualifiers
1..421
/organism="Mesembryanthemum crystallinum"
/db_xref="caxon:3544"
<1..>421
/gene="NHX2"
<1..>421
/gene="NHX2"
/codon_start=1
/product="putative Na+/H+-antiporter"
/protein_id="AAF91424.1"
/db_xref="GI:9652170"
/translaton="KKQFFRNFTLIVFGALSTLVSTIISLGAMEFFKKLDIGSLEL
GDYLAIGAFNATDSVCTLOVLHODEPPLYSIVFGGVNDATSVLVFNATQNFDT
HIDRIAFKFNGLYLFSTLGLATGLLSAYLIKK"
BASE COUNT 110 a 85 c 88 g 138 t
ORIGIN
alignment_scores:
Quality: 586.00 Length: 140
Ratio: 4.406 Gaps: 0
Percent Similarity: 95.000 Percent Identity: 80.000
alignment_block:
US-09-271-584A-2 x AF279671 ..
Align seg 1/1 to: AF279671 from: 1 to: 421
100 LysLysGlnPhePheArgAsnValThrIleMetLeuPheGlyAlaVal 116
|||||
1 AAGAGCAATTTTCCGGAATTCATTACAATCATAGTATTGGAGCCCT 50
|||||
116 lGlyThrIleSerCysThrIleIleSerLeuGlyValThrGlnPheP 133
|||||
51 GGGTACATTGGTGTCTATTACCATCATATACACTAGGAGCTAGGAGTCT 100
|||||
133 heLysLysLeuAspIleGlyThrPheAspLeuGlyAspTyrIleAlaIle 149
|||||
101 TCAGAGAGTTGGACATTGGTGTCTCTAGAGTAGGAGCTATCTTGCAATT 150
|||||
150 GlyAlaIlePheAlaLaThrAspSerValCysThrLeuGlnValLeuAs 166
|||||
151 GGTGCATATTCGCTGCACCGACTCTGTTGCACCTTCGAGGTGCTTCA 200
|||||
166 nGlnAspGluThrProLeuLeuTyrSerLeuValPheGlyGluGlyValV 183
|||||
201 TCAGGATGAGACCCCTCTACTATACAGTCTGGTGTGGGGAAGGTGAG 250
|||||
183 alAsnAspAlaThrSerValValPheAsnAlaIleGlnSerPheAsp 199
|||||
251 TCAATGATGACATCATAGTGGTCTTTTAAATCCCATCAGAACTTTGAC 300
|||||
200 LeuThrHisLeuAsnHisGluAlaAlaPheHisLeuLeuGlyAsnPheLe 216

```

seq\_name: gb\_pr5:D87743

seq\_documentation\_block:

LOCUS D87743 4408 bp mRNA PRI 10-JUL-1997

DEFINITION Human mRNA for KIAA0267 gene, partial cds.

ACCESSION D87743

VERSION D87743.1 GI:1665826

KEYWORDS KIAA0267.

SOURCE Homo sapiens male bone marrow Myeloblast cell\_line:KG-1 cDNA to mRNA, clone\_lib:library in pBluescript II SK clone:HA7045.

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 4408)

AUTHORS Nomura,N.

TITLE Direct Submission

JOURNAL Submitted (05-SEP-1996) to the DDBJ/EMBL/GenBank databases. Nobuo Nomura, Kazusa DNA Research Institute, Gene Structure 1; 1532-3 Yana, Kizarazu, Chiba 292, Japan (E-mail:cdnainfo@kazusa.or.jp, URL: http://www.kazusa.or.jp, Tel:0438-52-3930, Fax:0438-52-3931)

REFERENCE 2 (sites)

AUTHORS Nagase,T., Seki,N., Ishikawa,K., Ohara,O. and Nomura,N.

TITLE Prediction of the coding sequences of unidentified human genes.VI. The coding sequences of 80 new genes (KIAA0201-KIAA0280) deduced by analysis of cDNA clones from Human cell line KG-1 and brain

JOURNAL Unpublished (1996)

REFERENCE 3 (sites)

AUTHORS Nagase,T., Seki,N., Ishikawa,K., Ohira,M., Kawarabayashi,Y., Ohara,O., Tanaka,A., Kotani,H., Miyajima,N. and Nomura,N.

TITLE Prediction of the coding sequences of unidentified human genes.VI. The coding sequences of 80 new genes (KIAA0201-KIAA0280) deduced by analysis of cDNA clones from cell line KG-1 and brain

JOURNAL DNA Res. 3 (5), 321-329 (1996)

MEDLINE 97191544

FEATURES

Location/Qualifiers

1..4408

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/cell\_line="KG-1"

/cell\_type="Myeloblast"

/clone="HA7045"

/clone\_lib="Library in pBluescript II SK"

/sex="male"

/tissue\_type="bone marrow"

1..2001

/gene="KIAA0267"

<1..2001

/note="Similar to Human Na+/H+ exchanger 2 (A57644)"

/codon\_start=1

/protein\_id="BAAL3449.1"

/db\_xref="GI:1665827"

/translation="RGWRAPLRGVSSPPARLRPLMLLLAVGVDFWAGASDGGG

GEARAMDEEIVSEKQAEESHRQDSANLLITLTLTLWLFKRRARFLHETGLA

MTYGLLVGLVRYGHVPSDNNVTLSCEVQSSPTLLVTFDPEVFNILLPIFYA

GYSKKRRHFRNGLSILAYAPLGTATSCFVIGSIMYGVTLMKVTGLAGDFTDCL

LFGATVSATDPVTVAIFHELOVDVLYALLFGESLVNDAIVLSSSIIVAYOPAGDN

SHTDVTAMFKSIGIFIGISFGSFGAMGATGVVTLVTKTKRLREQLLETGFFLMS

WSTFLLAEAWGFTGVAVFCGIIQAHYTYNNLTSESQHRTKQLFELNLEAFNFI

YNGLFTFTQNHVFNPTVVGAFVAFILGRANILPLSLLNLRGRSKIGSNQHHMM

FAGLRGAMAFALNDTATYARQMMFTLLTIVFTVWFGGGTTAMLSCILIRVGD

SDQHLGVPENERERTTKAESAWLFMRMYNFDHNYLKPDLTHSGPPPTTLTPACCGPIA

222 erThrLeuLeuGlyAlaAlaThrGlyLeuIleSerAlaTyrValIleLys 230





|   |  |
|---|--|
| JOURNAL   | Submitted (28-OCT-1999) IBLs-Molecular Genetics, University of Glasgow, Glasgow, Scotland G11 6NU, UK  |
| FEATURES  | Location/Qualifiers  |
| source  | 1. 2604  |
|   | /organism="Drosophila melanogaster"  |
|   | /db_xref="taxon:7227"  |
|   | /chromosome="2"  |
|   | /map="2L; 26F5-27A2"   |
| gene  | 1. 2604  |
|   | /gene="NHE3"   |
| CDS   | 385.. 2448   |
|   | /gene="NHE3"   |
|   | /note="sodium-hydrogen exchanger"  |
|   | /codon_start=1   |
|   | /product="sodium-hydrogen exchanger NHE3"  |
|   | /protein_id="AAF13702.1"   |
|   | /db_xref="GI:6470267"  |
|   | /translation="MLLVGFALCVLAPGCSATDTDLADAKATLNHRTQSLDLVLFV<br>FLALTLTLIWLFKHRRVSNLHETGLAVIYGLIVAGIIRYAGTSGLTDHMOVPEQGV<br>TYSKLPDPTLWFRYPVNTGTGTPPEGIKYAYVFRQVHDENEIDLKATFDPEV<br>FENILPPIIFYAGYCLKKYFFRNLGAILTFAIVGTTLTSAFLIGGMYGCVKLMPIY<br>LSSPTFLDSLIFGALISPTDPLTILAFNDLRVDVNLVAILVIGESVLDVAIVLSG<br>ALQNGEHSYNTGEFETAFELSLSDPESIFLLSLMTGAAGCLTALMTKFTVRDPE<br>LLESAFLVMSITFLJAEATFELGVAVLFCGICQAHYYNNLSDSRSRQTKQIEL<br>LNLAENFIYSYIGVSMFTFPKHFDAGFIITAFICAAIGRAVNVYPLSLNLLNKKRP<br>KISTNFQHMLFFAGLRGAMSFALIRNTVSDERQTLTATSLIVFTVVIQGGAAFL<br>LNLWLIKPVGVDDTEQLNVOYHSSDGYLODEVGGGVNRRNMRSLSGTDTNLDTNP<br>TNGSISGASGGRRNSHEKALLARIWGNFDTKYMKPLLTSHSRPTLLFTLPVCCNPLAR<br>LLTTTQQLTQDGSERFRVDSDDICIDNDTNGLSQDGSNAAPGVGRNLSLRVGGD<br>H<br>RNAYI" |
| BASE COUNT                                      | 627 a 661 c 638 g 678 t  |
| ORIGIN  |  |
| alignment_scores:                               |  |
| Quality:  | 568.00 Length: 608   |
| Ratio:  | 1.656 Gaps: 23   |
| Percent Similarity:                             | 56.414 Percent Identity: 29.276  |
| alignment_block:                                |  |
| US-09-271-584A-2 x AF199463                     |  |
| Align seg 1/1 to: AF199463 from: 1 to: 2604     |  |
| 9   | LeuProSerLeuSerThrSerAspHisAlaSerValValAlaLeuAsnLe 25  |
|   | ::: :::     :::     ::: :::     :::     :::     :::  |
| 457   | CTGACGCCCAAGCCAGCTCAATCAC...CGCATCCAGAGCTCGACCT 503  |
| 25  | u.....PheValAlaLeuLeuCysAlaCysIleValLeuGlyHisLeuL 40   |
|   | ::: :::     :::     :::     :::  |
| 504   | TCTAGTCTCTGCTGCTCTGCTGGCGCTCACAGTGCACAAATTTGGCTTT 553  |
| 40  | euGluGluAsnArg.....TrpMetAsnGluSerIleThrAlaLeuLeu 54   |
| ::: :::     :::     :::     :::     :::     ::: |  |
| 554   | TCAAACACCATCGCGTCTCTGCTCCACGAGACGGGATTTGGCCGTCATC 603  |
| 55  | IleGlyLeuGlyThrGlyValThrIleLeuLeuIle...SerLysGlyLy 70  |
| :::     :::     :::     :::     :::             |  |
| 604   | TACGCTTGATTTGGGTGCAATATACGATATCCCGGACCTCCCGGCAC 653  |
| 70  | sSerSerHisLeuLeuVal..... 76  |
| :::     :::                                     |  |
| 654   | CCTGGACCATCATCGAGTGGAGCCTCAGGAGTTCCAAACGTACAGCGATA 703   |
| 76  | ..... 76   |
| 704   | AATTGCCGCTGACTCTTTTGGTTAGTACCCCGTTAATCAACAACAAAT 753   |
| 76  | ..... 76   |
| 754   | GGCAAGCGCCAGAGGAATCAAGACCATGTCCTACGTGTTTCGAGG 803  |
| 77  | .....Phes 78   |















Mon Mar 26 11:04:16 2001

```

1466 .....CACATCTCTCGCCATCGAGGACATATCCGGA.... 1498
491 euThrArgProThrArgThrValHisTyr.....TipArg 503
1499 .....CAGATCGGCGACAAATATCTCAGAGACAAGTGTCC 1534
504 GlnPheAspSerPheMetArgProVal 513
1535 CACTTCGACAGAGTCTCTCAGCAGGTC 1564
seq_name: /SID56/gcdata/geneseq/NA1993.DAT.Q51524
seq_documentation_block:
ID_Q51524 standard; DNA; 4151 BP.
XX
AC Q51524;
XX
DT 18-MAY-1994 (first entry)
XX
DE Rabbit sodium ion/hydrogen ion anti-transport carrier gene.
XX
KW Rabbit; sodium; hydrogen; anti-transport carrier; transgenic mouse;
XX model animal; hypertension; hypercardia; ss.
XX
OS Oryctolagus cuniculus.
XX
FH Key Location/Qualifiers
FT CDS 361..2811
FT FT /*tag= a
XX
PN JP05268856-A.
XX
PD 19-OCT-1993.
XX
PF 26-MAR-1992; 92JP-0068388.
XX
PR 26-MAR-1992; 92JP-0068388.
XX
PA (SUMU) SUMITOMO SEIYAKU KK.
XX
DR WPI: 1993-364207/46.
DR P-PSDB; R44143.
XX
PT Transgenic mouse expressing sodium ion-hydrogen ion
PT anti-transport carrier gene - useful for model animal of
PT hypertension and hypercardia
XX
PS Disclosure; Fig 8-12; 12pp; Japanese.
XX
CC Transgenic mice having this gene are useful as animal model for the
CC analysis of the effect of sodium ion/hydrogen-ion anti-transport
CC carrier gene activation on the regulation of blood pressure. The
CC transgenic animals are useful as models of hypertension and
CC hypercardia.
XX
SQ Sequence 4151 BP; 754 A; 1429 C; 1174 G; 794 T; 0 other;

alignment_scores:
Quality: 441.00 Length: 577
Ratio: 1.353 Gaps: 21
Percent Similarity: 56.499 Percent Identity: 25.303

alignment_block:
US-09-271-584A-2 x Q51524
Align seg 1/1 to: Q51524 from: 1 to: 4151
23 LeuAsnLeuPheValAlaLeuLeuCysAlaCysIleValLeuGlyHisLe 39
673 ATCTCCCTCTGGATCCTGCTGGCTGCCTCATGAAGATAGTTTCCACGT 722

```

```

723 GATCCCAACCATCTCCAGCATCGTCCGGAGAGCTGCCTCTGATGTGG 772
55 leGlyLeuGlyThrGlyValThrIleLeuLeuIleSerLysGlyLysSer 71
773 TGGGCTGCTGGTGGGGGCTGATC.....AAGGCGGTGGGC 810
72 SerHisLeuLeuValPheSerGluAspLeuPhePheIleTyrLeuLeuPr 88
811 GAGAAGCGCCCTTCTCGAGTCGAGGTCTTTTCTCTCTTCCTGCTGCC 860
88 opIleIlePheAsnAlaGlyPheGlnValLysLysLysGlnPhePheA 105
861 TCCCATCATCTCGGAGCGCGCTACTTCTCGCGCTCGCGCAGTTCACCG 910
105 rGAsnPheValThrIleMetLeuPheGlyAlaValGlyThrIle..... 119
911 AGAACCTGGGACCATCTCTATCTCGCCGTGGTGGCAGCGCTGTGAAC 960
120 .....IleSerCysThrIleIleSerLeuGlyValTh 130
961 GCCTTCTTCTCGGCGGCTCATGTACGCGGTGCTGTGGTGGCGCGGA 1010
130 rGlnPhePheLysLysLeuAspIleGlyThrPheAspLeuGlyAspTyrL 147
1011 GCAG.....ATCAACACATCGGCTCTCTGGACAACC 1042
147 euAlaIleGlyAlaIlePheAlaIleThrAspSerValCysThrLeuGln 163
1043 TGCTGTGGGACATCATCTCGCGGTGACCGCGTGGCGTGGCGTGGCG 1092
164 ValLeuAsnGlnAspGlu...ThrProLeuLeuTyrSerLeuValPheG 179
1093 GTCTTCGAGGAAATCCACATCAATGAGTGTGTCACATCTCTGGTTCGG 1142
179 yGluGlyValValAsnAspAlaThrSerValValValPheAsnAlaIleG 196
1143 CGAGTCCCTGCTCAAGGCGCGTCACTGTGTGTGTATCACCTCTTTG 1192
196 InSerPheAspLeuThrHisLeuAsnHisGluAlaAlaPheHisLeuLeu 212
1193 AGGAGTTT.....GCCAACTACGACCGTGGGCTCTCTGCTACAGCTAC 1236
213 GlyAsnPheLeuTyrLeuPheLeuLeuSer.....ThrLeuLeuG 226
1237 CTGGGCTTCTGAGCTTCTCTGTGTGGCGCTCGCGGGGGTGTCTGTCGG 1286
226 yAlaAlaThrGlyLeuIleSerAlaTyrValIleLysLysLeuTyrPheG 243
1287 CGTGGTCTACGGGTCTACGAGCCTTCACTCCCGA.....TTCA 1327
243 lyArgHisSerThrAspArgGluValAlaLeuMetMetLeuMetAlaTyr 259
1328 CCGCCACACATCCCGCTCATCGAGCGCGCTCTCTCTCTCTACAGCTAC 1377
260 LeuSerTyrMetLeuAlaGluLeuPheAspLeuSerGlyIleLeuThrVa 276
1378 ATGGCTTACCTGTGCGCGGAGCTCTTCCACCTGTGCGGATCATGGCACT 1427
276 lPhePheCysGlyIleValMetSerHisTyrThrTrpHisAsnValThrG 293
1428 CATCGCTCGGAGTGTGTATGCGGCTTACGTGGAGGCGGAACATCTCCG 1477
293 luSerSerArgIleThrThrLysHisThrPheAlaThrLeuSerPheLeu 309
1478 ACAAGTCGACACACCATCACTTCTCTGAAGATGTGGAGCGCGCTC 1527
310 AlaGluThrPheIlePheLeuTyrValGlyMetAspAlaLeu...Asp1 325
1528 AGCAGGACCTCATCTTCTCTCTCTCGCGCTCTCCACCGTGGCGGCTC 1577
325 eAsnLysTrpArgSerValSerAspThrProGlyThrSerIleAlaValS 342

```



```
1578 CCACCACTGGAATTGGACCTTCGAC.....ATCA 1606
342 erSerIleLeuMetGlyLeuValMetValGlyArgAlaAlaPheValPhe 358
|||||.....|
1607 GCACCTGCTC.....TTCGCTCATCGCCGCTGCTGGGTGCTG 1650
359 ProLeuSerPheLeuSerAsnLeuAlaLysLysAsnGlnSerGluLysI 375
|||||.....|
1651 GGCTGACCTGGTTCATCAAC.....AGTTCGCATAGTAGGCT 1691
375 eAsnPheAsnMetGlnValValIleTrpSerGlyLeuMetArgGlyA 392
|||||.....|
1692 GACGCCAAGACGACGATTCATCGCTTTCGAGGCGCTG.....CGGGCG 1738
392 laValSerMetAlaLeuAlaTyr.....AsnLysPheThrArg 404
|||||.....|
1739 CCATGCGCTTCGCTGGCTACTCTCTGGACAAAGACGACTTCCCATG 1788
405 AlaGlyHisThrAspValArgGlyAsnAlaIleMetIleThrSerThr 421
|||||.....|
1789 TGGCAC.....CTGTTCTCACCCTCATCAT 1814
421 eThrValCysLeuPheSerThrValValPheGlyMetLeuThrLysPro 438
|||||.....|
1815 CACCGTCATCTTCTCACCCTTTCGAGGCGCATGACATCCGCCCGC 1864
438 euIleSerTyrLeuLeuProHisGlnAsnAlaThrThrSerMetLeuSer 454
|||||.....|
1865 TGTGGACCTGCTGGCGCTGAAAGAAAGCAGGAAACAAAGCGC..... 1908
455 AspAspAsnThrProLysSerIleHisIleProLeuLeuAspGlnAsp 471
|||||.....|
1909 .....TCCATCAACGAGGAGATCCACACGCTTCTCTGAC..... 1944
471 rPheIleGluProSerGlyAsnHisAsnValProArgProAspSerIleA 488
|||||.....|
1945 .....CACCTTCTGACAGCAGCAGGAGCATCA 1972
488 rGlyPheLeuThrArgProThrArgThrValHisTyrTyrTrpArg... 503
|||||.....|
1973 GAGGCCACTACGCG.....CACCACCATTTGGAAGAC 2004
504 .....GlnPheAspAspSerPheMetArgProValPhe..... 514
2005 AGCTCAACCGGTTTACAAAGAGTACGTGAAGAGTGTCTGATCGCGG 2054
514 ..... 514
2055 CGAGCGCTCCAAGGAGCGCCAGCTCATCGCTTACCACAGATGGAGA 2104
515 .....GlyGlyArgGlyPheValPro 521
2105 TGAAGCAGGCGCATCGAGCTGGTGGAGAGTGGGGCATGGGCAAGATCCC 2154
522 PheValProGlySer.....Pr 527
2155 TCGGCGCTCTCCACCGCTCTCCATGCGACATCCACCCCAAGCGCTGCC 2204
527 oThrGluArgAsnProProAspLeuSerLys 537
|||||.....|
2205 TGGCAGCGCATCTCGCGCGCTGTCCAG 2235
```

seq\_name: /SID56/legdata/geneseq/geneseq/NA1999.DAT:22611

seq\_documentation\_block:

ID 22611 standard; DNA; 330 BP.

XX  
AC 22611;XX  
DT 18-JAN-2000 (first entry)XX  
DE Arabidopsis Na/H antiporter transporter gene from EST clone T75860.

```
KW Sodium, proton; antiport; transporter; salt tolerance; salt management;  
KW transgenic plant; survival; soil; farming; accumulation; irrigation;  
KW crop; ss  
XX  
OS Arabidopsis thaliana.  
XX  
PN W09947679-A2.  
XX  
PD 23-SEP-1999.  
XX  
PF 18-MAR-1999; 99MO-CA00219.  
XX  
PR 18-MAR-1998; 98US-0078474.  
PR 15-JAN-1999; 99US-0116111.  
XX  
PA (BLUM/) BLUMWALD E.  
PA (APSE/) APSE M.  
PA (SNED/) SNEDDEN W.  
PA (AHAR/) AHARON G.  
XX  
PI Blumwald E, Apse M, Snedden W, Aharon G;  
XX  
DR WPI; 1999-571840/48.  
XX  
PT Nucleic acid molecules encoding sodium/proton transport polypeptides,  
XX useful in genetic engineering salt tolerance in crop plants  
XX Disclosure; Fig 8H; 9pp; English.  
XX  
CC The invention relates to an isolated nucleic acid molecule encoding  
CC a plant Na/H antiport (PNHX) transporter polypeptide, or a fragment  
CC and capable of increasing salt tolerance in a cell. This sequence  
CC corresponds to an Na/H antiport transporter gene fragment from the  
CC Arabidopsis expressed sequence tag (EST) library clone T75860. The  
CC Na/H transporter polypeptides provide a means of intracellular salt  
CC management, particularly in plants. The sequences are useful for  
CC producing transgenic plants that are capable of surviving in soil  
CC with high salt levels that would normally inhibit growth of the crop  
CC species. This would be useful in farming land in areas that are generally  
CC considered unproductive through salt accumulation and poor irrigation,  
CC e.g. in India, Australia, and prairies in USA or Canada. Commercial  
CC crops, such as potatoes, tomatoes, brassica, cotton, sunflower,  
CC strawberries, spinach, lettuce, rice, soybean, corn, wheat, barley,  
CC atriplex, sorghum, alfalfa, salt cornia and others would benefit from  
CC increased salt tolerance.  
XX  
SQ Sequence 330 BP; 92 A; 68 C; 66 G; 91 T; 13 other;
```

alignment\_scores:

Quality: 415.00

Ratio: 5.061

Percent Similarity: 96.471 Percent Identity: 95.294

alignment\_block:

US-09-271-584A-2 x 22611

Align seg 1/1 to: 22611 from: 1 to: 330

```
358 PheProLeuSerPheLeuSerAsnLeuAlaLysLysAsnGlnSerGluLys 374  
|||||.....|  
1 TTTCCGTTATCGTTCTATCTAACTTAGCCAAAGAAATCAAGCGAGAA 50  
374 sIleAsnPheAsnMetGlnValValIleTrpSerGlyLeuMetArg 391  
|||||.....|  
51 AATCACTTTAATCATGAGGTGATTTGGTGTCTGCTCATGAGAG 100  
391 lYAlaValSerMetAlaLeuAlaTyrAsnLysPheThrArgAlaGlyHis 407  
|||||.....|  
101 GTGCTGTATCTATGCTCTTCATACACAGTTTACAAAGGCCGCGCAC 150  
408 ThrAspValArgGlyAsnAlaIleMetIleThrSerThrIleThrValCy 424  
|||||.....|
```



```
1156 .....CACTTGATG.....TGTATTATTT.....1176
224 uueGlyAlaAlaThrGlyLeuIleSerAla.....TyrValIleL 238
1177 .CATGGCTGGGATTCACATGATTCCTGCTGTATTACACGCTGTCCA 1225
238 ySLyLeuTyrPheGlyArgHis.....SerThrAspArgGluVal 251
1226 GAAGATATATACATACACACTGCTGGATCAGTCTGTGATACCAACTC 1275
252 AlaLeuMetLeuMetAlaTyrLeuSerTyrMetLeuAlaGluLeuPh 268
1276 CTCATCATATCATGCGCCCAATTTGCTGCTTTATTTGGTGAATCTTT 1325
268 eAspLeuSerGlyIleLeuThrValPhePheCysGlyIleValMetSerH 285
1326 TTTCCTATTATATTTGACGTTCTTATCACCAGTTAAAGTTACTC 1375
285 lSyrThrTrpHisAsnValThrGluSerSerArgIleThrThrLysHis 301
1376 ACCAAGCAGATCTAATCTGTACATGAAGCT.....GTGAGAGCT 1416
302 ThrPheAlaThrLeuSerPheLeuAlaGluThrPheIlePheLeuTyrVa 318
1417 ACCTTTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1461
318 lGlyMetAspAlaLeuAspIleAspLysTrpArgSerValSerAspThrP 335
1462 .....CCATGGCGA.....C 1471
335 roGlyThrSerIleAla.....ValSerSerIleLeu 345
1472 CTGAAGGAAGATTCGAGAAGATATATGATTACATCAATAACATCTCTC 1521
346 Met.....GlyLeuValMetValGlyArgAlaAlaPheValPhePr 359
1522 ATGCATCATCAGGCTCTCTGGTA.....TCTACAAATTTCTGCTTCTT 1565
359 oLeuSerPheLeuSerAsnLeuAlaLysLysAsnGlnSerGlu...LysI 375
1566 TATGAGAGGTTCAAGCAATCTGAGAAGAACTGGAATCAATATAAAA 1615
375 leAsnPheAsnMetGlnValValIleTrpSerGlyLeuMetArgGly 391
1616 TCCAATTTGGA..... 1626
392 AlaValSerMetAlaLeuAlaTyrAsnLysPheThrAlaAlaGlyHisTh 408
1627 .....ACAACTTTTCC.....CACTC 1643
408 rAspValArgGlyAsnAlaIleMetIleThrSerThrIleThr 422
1644 AGATACTCTCGTAGTCATCTTACAGGTTTCAACAATCAGC1686

seq_name: /cn2_6/ptodata/1/ina/5A_COMB.seq:us-08-098-141-1
seq_documentation_block:
; Sequence 1, Application US/08098141
; Patent No. 5441875
; GENERAL INFORMATION:
; APPLICANT: Hediger Ph.D., Matthias A.
; TITLE OF INVENTION: Urea Transporter Polypeptide
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Choate, Hall & Stewart
; STREET: Exchange Place, 53 State Street
; CITY: Boston
; STATE: MA
; COUNTRY: U.S.A.
; ZIP: 02119
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
```

```
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/098.141
FILING DATE: 19930723
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Kaplan Esq., Warren A.
REGISTRATION NUMBER: 34,199
REFERENCE/DOCKET NUMBER: 092662-001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-227-5020
TELEFAX: 617-227-7566
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 3060 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-098-141-1

alignment_scores:
Quality: 98.50 Length: 319
Ratio: 0.724 Gaps: 17
Percent Similarity: 42.633 Percent Identity: 21.317

alignment_block:
US-09-271-584A-2 x US-08-098-141-1
Align seg 1/1 to: US-08-098-141-1 from: 1 to: 3060

34 IleValLeuGlyHisLeuLeuGluGluAsnArgTrpMetAsnGluSerIl 50
1262 ATCTGTATGGC...CTCTTTGTCCAGAACCCATGGTGG.....GCCAT 1302
50 eThrAlaLeuLeuIleGlyLeuGlyThrGlyValThrIleLeuLeuL 67
1303 CGCGGGATGCTGGGACCGCTCATGTCCACCTTGACTGCCCTCATCTGA 1352
67 erLysGlyLysSerSer.....His..... 73
1353 GCCAGGACAGGTACGCCATCCCTCCGACTCCACGGCTACACGGCGTG 1402
74 .....LeuLeuValPheSerGlu.....AspLeuPhePh 83
1403 CTGGTGGGCTGCTGATAGCGGTGTCTCCGACAAAGGCGGACTATTACTG 1452
83 eIleTyrLeuLeuProPheIleIlePheAsnAlaGlyPheGlnValLysL 100
1453 GTGGCTTCTGCTCCCTCTCATCGTGTATGTC..... 1483
100 ySLysGlnPhePheArgAsnPheValThrIleMetLeuPheGlyAlaVal 116
1483 ..... 1483
117 GlyThrIleIleSerCysThrIleIleSerLeuGlyValThrGlnPhePh 133
1484 .....ATGCTCTCCCATCTCTCCAGTGCCCTGGGCGCATCTT 1524
133 eLysLysLeuAspIleGlyThrPheAspLeuGlyAspTyrLeuAlaIleG 150
1525 CAGCAAGTGGGACCTCCAGCTTCTCACACTGCCCTTCAACATCGCCGTGA 1574
150 lYAlaIlePheAlaAlaThrAspSerValCysThrLeuGlnValLeuAsn 166
1575 CCCTGTACTTGGCAGCCACAGGCCAC..... 1600
167 GlnAspGluThrProLeuLeuTyrSerLeuValPheGlyGluValVa 183
```









```

114 GlyAlaValGlyThrIleSerCysThrile.....IleSe 126
...
334 GCTATCATCGGAACCTTTGCGTTCAGTGTGTACTGGTGGTCTGGTTA 383
126 rLeuGlyValThrGlnPhePheLeuLysLeuAspIleGlyThrPheAspL 143
126 rLeuGlyValThrGlnPhePheLeuLysLeuAspIleGlyThrPheAspL 143
384 TCTTGGCGCTCTATGATCTC.....ATGTATAAAGCTCCCT 421
143 euGlyAspTyrLeuAlaIleGlyAlaIlePheAlaAlaThrAspSerVal 159
...
422 TTGTTGAGTCTTATGTTGGTGGTCACTATATACGCTACGACCGCTGTC 471
160 CysThrLeuGlnValLeuAsnGlnAspGluThrProLeu...LeuTyrSe 175
472 ACTGTACTCTATATATTCAGGATGCGGACCGATGTTAAACCTCTATGC 521
175 rLeuValPheGlyGluGlyValValAsnAspAlaThrSerValValValp 192
...
522 TTGTTGCTTTGGGAATCAGTCTCTGATGCTATGCTATGCTATGCTATG 571
192 heAsnAlaIleGlnSerPheAspLeuThrHisLeuAsnHisGluAlaAla 208
572 ACAGAAATGTCCTTAGTAAACCCGACGCTCTGCTGCGGGAACATTTT 621
209 PheHisLeuGlyAsnPheLeuTyrLeuPheLeuLeuSerThrLeuLe 225
622 TTCATGGTGGTATCAGGTTTTCAGACTTTTGGTGGCTCAATGCTGTC 671
225 uGlyAlaAlaThrGlyLeuLeuSerAlaTyrValIleLysLysLeuTyrp 242
672 AGGGTTGGGTTGGATTACCTCAGCTTACTCTTTAAG.....TAGT 715
242 heGly.....ArgHisSerThrAspArgGluValAlaLeuMetMet 255
716 CAGGATTGGACACCGAGATCTTCAGAACTTGGAGTGTTCCTCTTTGTA 765
256 LeuMetAlaTyrLeuSerTyrMetLeuAlaGluLeuPheAspLeuSerGl 272
766 CTTTCCCGTATTTTCATACATGCTTCAGAGAGGTGTTGGTCTCTCCGG 815
272 yLeLeuThrValPhePheCysGlyIleValMetSerHisTyrThrTrph 289
816 CATTTGTTCTATCTCTTCACAGGAATGTTATGAAGCGCTACACTTCT 865
289 isAsnValThrGluSerSerArgIleThrLysHisThrPheAlaThr 305
866 CAATCTCTCAGAGCTTCACAGAGTTCGTATCTCTTTTTCCTTGG 915
306 LeuSerPheLeuAlaGluThrPheIlePheLeuTyrValGlyMetAsp.. 321
916 ATATCTCGTAGCAGAACTTTCAGTTCATTTACATGGGATTTGATAT 965
322 .AlaLeuAspIleAspLysTyrArgSerValSerAspThrProGlyThrS 338
966 TGGCATTGGAGCAGCATAGTGGTCCCAT..... 993
338 erIleAlaValSerSerIleLeuMetGlyLeuVal..... 349
994 .....GTTGGTTATCTCTTCTCTATGTTATCTCTATCTCTATCTGAT 1035
350 .....MetValGlyArgAlaAlaPheValPheProLeuSerPheLe 363
1036 CGTCAGTGTATGATGTCAGTGGCTCAATGTTATTTGGTGGTGCATATTT 1085
363 uSerAsnLeuAlaLysAsnGlnSerGluLysIleAsnPheAsnMetG 380
1086 GGTCAACCTATTAGACAG...GAGAACCAAGAGATACCTATGAGCAC 1132
380 InValValIleThrPheSerGlyLeuMetArgGlyAlaValSerMetAla 396
1133 AAAAGCCCTTTGGTATAGTGGACTT...CGAGGGCAATGGCATTTGCA 1179

```

```

397 LeuAla.TyrAsnLysPheThrArgAlaGlyHisThrAspValArgGlyA 413
1180 CTTGCACCTTCATCATCATGATCATCAGAGGGTCACGCCCAATCAT 1229
413 snAlaIleMetIleThrSerThrIleThrValCysLeuPheSerThrVal 429
1230 CT.....TTACTGCAACCACTATTTGTTGTTCACGGTTTA 1270
430 ValPheGlyMetLeuThrLysProLeuIleSerTyrLeuLeuProHisGl 446
1271 CTAATAGGAGGTCGACAGGTAATAATGTTGGAAGCTTTG..... 1309
446 nAsnAlaThrThrSerMetLeuSerAspAspAsnThrProLysSerIleH 463
1310 .....GAAAGTTGAGTGACGAT..... 1327
463 isIleProLeuLeuAspGlnAspSerPhe.....IleGluProSer 476
1328 .....CTTGAT...GACTCCATGCTGAAGGCTTTGAAGAGAGC 1363
477 GlyAsnHisAsnValProArgProAspSerIleArgGlyPheLeuThrAr 493
1364 GATCATCATGATGTCCTCTCTCTTTAGCATTTGGAGCTTCATCTGACGA 1413
493 gProThr.....ArgThrValH 499
1414 GGATACATCATCATCAGAGAGGTTCAAGATCAAGCTGAAGGAGTTTC 1463
499 isTyrTyrTrpArgGlnPhe.....AspAspSerPheMetArgPro 512
1464 ACAAAACCACTACATCATTCACCGCGTTGGACAAAACCTTCTGACTCCG 1513
513 ValPhe 514
1514 TTCCTC 1519

seq_name: /SIDS6/gcgdata/geneseq/geneseqn/NA1997.DAT:T51330
seq_documentation_block:
ID T51330 standard; CDNA; 2584 BP..
XX
AC T51330;
XX
DT 29-APR-1997 (first entry)
XX
DE Human Na+/H+ exchanger isoform NHE3 composite cDNA.
XX
KW NHE3; Na+/H+ exchanger; sodium ion-proton exchanger; antiporter;
KW chronic metabolic acidosis; chronic renal failure; hypertension;
KW diabetic nephropathy; familial diarrhoea; cystic fibrosis; kidney;
KW small intestine; ss.
XX
OS Homo sapiens.
XX
FH Key
FH CDS 11..2515
FH /tag= a
FH /transl_except= pos:1970..1972..aa:Arg
FH /note= "codon CAC (bases 1870-1972) codes for Arg"
XX
PN WO9703196-A1.
XX
PD 30-JAN-1997.
XX
PF 10-JUL-1996; 96WO-US11670.
XX
PR 11-JUL-1995; 95US-0001061.
XX
PA (MAGA-) MAGAININ PHARM INC.
XX
PI Brant SR, Donowitz M, Tse C, Yun CHC;
XX
DR WPI; 1997-119051/11.

```



Downloaded from <http://ajphaphysocpharm.sagepub.com/> at 11:51 11 November 2014







XX Arabidopsis Na/H antiporter transporter gene from EST clone T75860.





```

1156 .....CACTGAG.....TGGTATTATTTT..... 1176
224 uNeuGlyAlaAlaThrGlyLeuSerAla.....TyrValIleL 238
1177 CATTGGCTGGGATTTCCATTGATTCCTGTTATTCACGCTGTGCCA 1225
238 yLyLeuTyrPheGlyArgHis.....SerThrAspArgGluVal 251
1226 GAAGATTATATACATGACCACTGCTGATCAGTCTGTGATACCACTC 1275
252 AlaLeuMetMetLeuMetAlaTyrLeuSerTyrMetLeuAlaGluLeuPh 268
1276 CTCTACATATCCATGGCCCAATTTGCTGCTTTTATTGTTGAATCTTT 1325
268 eAspLeuSeuGlyIleLeuThrValPhePheCysGlyIleValMetSerH 285
1326 TTTCTATTATATATGTTAGTCTGTTCTTATACCAAGTTAAAGTTACTC 1375
285 IsTyrThrTrpHisAsnValThrGluSerSerArgIleThrThrLysHis 301
1376 ACCAAGCAGAAATCAATCTGTATCAAGAAAGCT.....GTGAGAGCT 1416
302 ThrPheAlaThrLeuSerPheLeuAlaGluThrPheIlePheLeuTyrVa 318
1417 ACGTTATCTCTGGTGGCTGTGCTTGGCATTTGTTGCTGATT..... 1461
318 lGlyMetAspAlaLeuAspLeuAspLysTrpArgSerValSerAspThrp 335
1462 .....CCATGGCGA.....C 1471
335 roGlyThrSerIleAla.....ValSerSerIleLeu 345
1472 CTGAGGAGAGATTCAGAGAGATATATGATTACATCAATCAATCTCTC 1521
346 Met.....GlyLeuValMetValGlyArgAlaAlaPheValPhePr 359
1522 ATGCACTACTCAGGGTCTCTGGTA.....TCTCAATTTTCTGCTCTT 1565
359 oLeuSerPheLeuSerAsnLeuAlaLysLysAsnGlnSerGlu...LysI 375
1566 TAATGGAGAGGTTCAAGCAATCTCAGAGAACTGGGAATCAATATAA 1615
375 leAsnPheAsnMetGlnValValIleTrpTrpSerGlyLeuMetArgGly 391
1616 TCCAATTGGA.....AACAACTTTTCC.....CACTC 1643
392 AlaValSerMetAlaLeuAlaTyrAsnLysPheThrArgAlaGlyHisTh 408
1627 ..... 1627
408 AspValArgGlyAsnAlaIleMetIleThrSerThrIleThr 422
1644 AGATACCTCCGTGAGTGCATTCACACGGTTTCAACATCAGC 1686

```

seq\_name: /cgn2\_6/ptodata/1/ins/5A\_COMB.seq:us-08-098-141-1

seq\_documentation\_block:

; Sequence 1, Application US/08098141

; Patent No. 5441875

; GENERAL INFORMATION:

; APPLICANT: Hediger Ph.D., Matthias A.

; TITLE OF INVENTION: Urea Transporter Polypeptide

; NUMBER OF SEQUENCES: 2

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Choate, Hall & Stewart

; STREET: Exchange Place, 53 State Street

; CITY: Boston

; STATE: MA

; COUNTRY: U.S.A.

; ZIP: 02119

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

```

; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/098,141
; FILING DATE: 19930723
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Kaplan Esq., Warren A.
; REGISTRATION NUMBER: 34,199
; REFERENCE/DOCKET NUMBER: 092662-001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617:227-5020
; TELEFAX: 617:227-7566
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3060 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; US-08-098-141-1

```

# alignment\_scores:

```

Quality: 98.50 Length: 319
Ratio: 0.724 Gaps: 17
Percent Similarity: 42.633 Percent Identity: 21.317

```

# alignment\_block:

US-09-271-584A-2 x US-08-098-141-1

Align seg 1/1 to: US-08-098-141-1 from: 1 to: 3060

```

34 IleValLeuGlyHisLeuLeuGluAsnArgTrpMetAsnGluSerI 50
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1262 ATCGTGATTGGC...CTCTTTGTCAGAACCCATGTTGG.....GCCAT 1302
50 ePheAlaLeuLeuIleGlyLeuGlyThrGlyValThrIleLeuLeu 67
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1303 CGCGGGATGCCTGGCACCGTCATGTGCACCTTGACCTGCCTCATCTGA 1352
67 erLysGlyLysSerSer.....His..... 73
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1353 GCCAGGACAGGTACGATCGCTCCGCGACTCCACGGCTACAACGGCGTG 1402
74 .....LeuLeuValPheSerGlu.....AspLeuPhePh 83
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1403 CTGGTGGGCTGCTGATAGCGGTGTTCTCCGACAGGGCGACTATTACTG 1452
83 eIleTyrLeuLeuProPheIlePheAsnAlaGlyPheGlnValLysL 100
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1453 GTGGCTTCTGCTCCCTGTCATCGTGTGATGCC..... 1483
100 yLysGlnPhePheArgAsnPheValThrIleMetLeuPheGlyAlaVal 116
1483 ..... 1483
117 GlyThrIleIleSerCysThrIleIleSerLeuGlyValThrGlnPhePh 133
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1484 .....ATGCTCTGCCCATCTCTCTCCAGTGCCTGGGCGACCATCTT 1524
133 eLysLysLeuAspIleGlyThrPheAspLeuGlyAspTyrLeuAlaIleG 150
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1525 CAGCAAGTGGGACCTCCCGAGTCTTCACACTGCCCTTCAACATCGCGTGA 1574
150 lyAlaIlePheAlaThrAspSerValCysThrLeuGlnValLeuAsn 166
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1575 CCTGTACTTGGCAGCCACAGGCCAC..... 1600
167 GlnAspGluThrProLeuLeuTyrSerLeuValPheGlyGluGlyValVa 183

```



```
1601 .....TACACCTCTCTCCACACACTGCT 1629
183 lAsnAspAlaThrSerVal.....ValValPheAsnAlaIleGlnSerP 198
1630 GCAGCGCTGATCTCCGTCGCCACATCAGCTGTCAGAGATCCAG.... 1675
198 heAspLeuThrHisLeuAsnHisGluAlaAlaPheHisLeu..... 212
1676 .....GTGCCCTGCTCTGTAGA 1693
212 ..... 212
1694 GCCATCCCGTGGGAATGCGCCAGGTATGGCTGTGATACCCCTGGAC 1743
213 ....GlyAsnPheLeuTyr...LeuPheLeuLeuSerThrLeu..... 224
1744 TGGAGGCATTTTCTCTGCTGCTGTATATCTTACCTCTTATTGCT 1793
225 .....LeuGlyAlaAlaThrClyLeuIleSerAlaTyrValIle 237
1794 TGCATGCGGCAATGGATGCCACCATGGGATGTACGACACTCAGGATT 1843
238 .....LysLeuLeuTyrPheGlyArgHisSerThrAspArgG1 250
1844 GGCACACCCCTTTGACTCCATCTATTGGCTGTGTGCTTCAACAGCAC 1893
250 uValAla.....LeuMetMetLeuMetAlaTyrLeuS 261
1894 GTGCGCATGCAATGCTGCGGAGCATGTTCTACGTATCACCCTGGCAGA 1943
261 erTyrMetLeuAla.....GluLeuPhe..... 268
1944 CTCACCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1993
269 AspLeuSerGlyIleLeuThrValPhe..... 277
1994 GCCCTGACCAAGTATTATCTGTGTTGGATTACCAACCTGCACCTGCC 2043
278 PheCys 279
2044 CTTCTGC 2050
```

seq\_name: /cgn2\_6/ptodata/1/ina/6\_COMB.seq:US-08-602-809-3

seq\_documentation\_block:

; Sequence 3, Application US/08602809

; Patent No. 6008012

; GENERAL INFORMATION:

; APPLICANT: BERGSMAN, DEREK

; TITLE OF INVENTION: HUMAN SOMATOSTATIN-LIKE R

; NUMBER OF SEQUENCES: 3

; CORRESPONDENCE ADDRESS:

; STREET: P.O. Box 980

; CITY: Valley Forge

; STATE: PA

; COUNTRY: US

; ZIP: 19482-0980

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette

; COMPUTER: IBM Compatible

; OPERATING SYSTEM: DOS

; SOFTWARE: FastSeq for Windows Version 2.0

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/602,809

; FILING DATE: 13-JUN-1997

; CLASSIFICATION: 536

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: PCT/US95/16472

; FILING DATE: 15-DEC-1995

```
; APPLICATION NUMBER: US 08/357,675
; FILING DATE: 16-DEC-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Prestia, Paul F
; REGISTRATION NUMBER: 23,031
; REFERENCE/DOCKET NUMBER: P50277
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 601-407-0700
; TELEFAX: 610-407-0701
; TELEX: 846169
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1080 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; US-08-602-809-3
```

alignment\_scores:

Quality: 96.50 Length: 371

Ratio: 0.652 Gaps: 22

Percent Similarity: 39.892 Percent Identity: 22.642

alignment\_block:

US-09-271-584A-2 x US-08-602-809-3 ..

Align seg 1/1 to: US-08-602-809-3 from: 1 to: 1080

```
171 ProLeuTyrSerLeuValPheGlyGluGlyValValAsnAspAlaTh 187
||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
94 CCTACCTCTACAGCATCATCTTTGCTGGTGGGAATA..... 129
187 rSerValValPheAsnAlaIleGlnSerPheAspLeuThrHisLeuA 204
129 ..... 129
204 snHisGluAlaAlaPheHisLeuLeuGlyAsnPheLeuTyrLeuPhe 220
||| ||||| ||||| ||||| ||||| ||||| |||||
130 .....TTTGAACACAGCTTTGGTGGTCAATCTC 156
221 LeuSerThrLeuLeuGlyAlaAlaThrClyLeuIleSerAlaTyrVal 237
||| ||||| ||||| ||||| ||||| ||||| |||||
157 ATT.....TACTTTTA 167
237 eLysLysLeuTyrPheGlyArgHisSerThrAspArgGluValAlaLeu 254
||| ||||| ||||| ||||| ||||| ||||| |||||
168 CATGAAGCTG.....AAGACTGTGGCCGCG 193
254 etMetLeuMetAlaTyrLeuSerTyrMetLeuAlaGluLeuPheAspLeu 270
||| ||||| ||||| ||||| ||||| ||||| |||||
194 TCTTTCTT.....CTCAATCTCGCTTGGCT.....GACTTA 225
271 SerGlyIleLeuThrValPhePheCysGlyIleValMetSer...His 286
||| ||||| ||||| ||||| ||||| ||||| |||||
226 TGCTTTTCTGACTTTCCTGCTGGGAGCTCTATACCCCTATGGAGTA 275
286 rThrTrpHisAsnValThrGluSerSerArgIleThrThrLysHisThrP 303
||| ||||| ||||| ||||| ||||| ||||| |||||
276 CGCTGGCCCTTCGGCAATCACCCTATATAGATCGCTTCG..... 315
303 heAlaThrLeuSerPheLeuAlaGluThrPheIlePheLeuTyrValGly 319
||| ||||| ||||| ||||| ||||| ||||| |||||
316 ..GCCAGCTGAGCTTCAACCTCTACCCAGCTGCTGTTCTCT..... 354
320 MetAspAlaLeuAspIleAspLysTrpArgSerValSerAspThrProL 336
||| ||||| ||||| ||||| ||||| ||||| |||||
355 CTCACGTGTCACGATCAGCCGCTAC..... 381
336 yThrSerIleAlaValSerSerIleLeuMetGlyLeuValMetValGlyA 353
381 ..... 381
```

